BELL/SLOMSON MODELS AND ULTRAPRODUCTS

Models and Ultraproducts:

AN INTRODUCTION

J. L. BELL

A. B. SLOMSON

511.
3
BEL

NORTH-HOLLAND / AMERICAN ELSEVIER

MODELS AND ULTRAPRODUCTS
AN INTRODUCTION

PRESTON POLYTECHNIC
LIBRARY & LEARNING RESOURCES SERVICE
This book must be returned on or before the date last stamped

12. MAR 1976
30. JUN. 1978
29. SEP. 1978
15. DEC. 1978
14. JUL 1989
18. OCT. 1991
30. NOV. 1992

110143
BELL, J.L.
Models and ultraproducts: an introduction.

110143

Chor
Lanc
Mars
Poul
Prest

511.3 BEL
A/C 110143
30107 000 526 159

Models and Ultraproducts:

AN INTRODUCTION

---

by

J. L. BELL

London School of Economics

and

A. B. SLOMSON

University of Leeds

Third revised printing

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM · OXFORD
AMERICAN ELSEVIER PUBLISHING COMPANY, INC. – NEW YORK

© North-Holland Publishing Company – 1969

All rights reserved. No part of this publication may be reproduced, stored in a retrieval systeem, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

Library of Congress Catalog Card Number: 78-87273
North-Holland ISBN: 0 7204 2054 7
American Elsevier ISBN: 0 444 10615 4

Published by:
North-Holland Publishing Company – Amsterdam
North-Holland Publishing Company Ltd. – Oxford

Sole distributors for the U.S.A. and Canada:
American Elsevier Publishing Company, Inc.
52 Vanderbilt Avenue
New York, N.Y. 10016

1st edition 1969
2nd revised printing 1971
3rd revised printing 1974

ACCESSION No. 110143
CLASS No.
511.3 BEL
23 NOV 1977
C/S | N | CATEGORY
| √ | NORMAL

Printed in The Netherlands

There is only one truth value: the truth
LUCIANO BÜCHLER

There is no book so bad that there is not something good in it
MIGUEL DE CERVANTES

The only rational way of educating is to be an example –
of what to avoid, if one can't be the other sort
ALBERT EINSTEIN

CONTENTS

INTRODUCTION 1

Ch. 0 PREREQUISITES 4

Ch. 1 BOOLEAN ALGEBRAS 7

1. Lattices 7 2. Filters 12 3. Ultrafilters 14
4. Homomorphisms and quotient algebras 17 5. Filters in topology 22 6. The Stone representation theorem 23
7. Atomless Boolean algebras and Cantor spaces 27
8. Historical and bibliographical remarks 31

Ch. 2 PROPOSITIONAL CALCULUS 32

1. The system SC 32 2. The Lindenbaum algebra 40
3. The completeness of propositional calculus 42 4. The compactness of propositional calculus 44 5. Historical and bibliographical remarks 49

Ch. 3 PREDICATE CALCULUS 50

1. The language L 50 2. Interpretations of L 54
3. An axiom system for predicate calculus 57 4. The Lindenbaum algebra of PC 61 5. The completeness of PC 62
6. Uncountable languages 67 7. Constants 68
8. Function symbols 69 9. Historical and bibliographical remarks 71

Ch. 4 BEGINNING MODEL THEORY 72

1. The basic notions 72 2. Unions of chains 79
3. Löwenheim-Skolem theorems 80 4. Hanf numbers 84
5. Historical and bibliographical remarks 86

Ch. 5 INTRODUCING ULTRAPRODUCTS 87

1. The ultraproduct construction 87 2. Łoš's theorem 89
3. Finite axiomatizability 92 4. The compactness theorem 102
5. The completeness theorem, the axiom of choice and the ultrafilter theorem 103 6. Historical and bibliographical remarks 106

Ch. 6 SET THEORETIC PROPERTIES OF ULTRAPRODUCTS 107

1. More about ultrafilters 107 2. Set theoretic properties of ultraproducts 122 3. The cardinality of ultraproducts 125
4. Well-ordered ultraproducts 133 5. The Rabin–Keisler theorem 136 6. Historical and bibliographical remarks 139

Ch. 7 ELEMENTARY CLASSES 140

1. The various sorts of elementary classes 140 2. Keisler's ultrapower theorem 145 3. Characterizing elementary classes 151 4. Craig's interpolation lemma 153
5. EC as a Boolean algebra 157 6. Historical and bibliographical remarks 159

Ch. 8 ULTRALIMITS 161

1. Two basic lemmas 161 2. Ultralimits 165
3. Elementary classes and ultralimits 169 4. Ultralimits and Craig's interpolation lemma 170 5. Historical and bibliographical remarks 174

Ch. 9 COMPLETENESS AND MODEL COMPLETENESS 175

1. Completeness 175 2. Model completeness 183
3. Universal and existential sentences 185 4. A test for model completeness 188 5. Applications of Robinson's test 191
6. Historical and bibliographical remarks 200

Ch. 10 HOMOGENEOUS AND UNIVERSAL STRUCTURES 201

1. The main definitions 201 2. The uniqueness of $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures 204 3. The existence of $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures 4. Full expansions 213 5. Historical and bibliographical remarks 216

Ch. 11 SATURATED MODELS 217

1. Saturated models 217 2. Ultraproducts and saturated structures 222 3. More about saturated structures 224 4. Special relational structures 229 5. Historical and bibliographical remarks 232

Ch. 12 APPLICATIONS OF ULTRAPRODUCTS 233

1. Gödel's completeness proof 233 2. A non-standard model of arithmetic 236 3. MacDowell and Specker's theorem 240 4. Cardinal-like orderings 244 5. Two-cardinal theorems 245 6. Vaught's two-cardinal theorem 249 7. Chang's two-cardinal theorem 253 8. Historical and bibliographical remarks 259

Ch. 13 GENERALIZED QUANTIFIERS 260

1. Generalized quantifiers 261 2. The quantifiers Q_α 262 3. The Chang quantifier 267 4. Fuhrken's reduction technique 270 5. Löwenheim-Skolem and compactness theorems for L_Q 275 6. Completeness theorem for L_Q 279 7. Historical and bibliographical remarks 285

Ch. 14 INFINITARY LANGUAGES 286

1. The language $L_{\alpha\beta}$ 286 2. The compactness property; incompactness results for accessible cardinals 289 3. Incompactness results for inaccessible cardinals 293 4. Measurable cardinals 295 5. Keisler's ultrapower proof 298 6. Measurable cardinals and the axiom of constructibility 300 7. Historical and bibliographical remarks 305

SOME SUGGESTIONS FOR FURTHER READING 308

BIBLIOGRAPHY 309

INDEX OF NAMED THEOREMS 317

GENERAL INDEX 319

PREFACE TO THE SECOND PRINTING

This second printing has been photographically produced from the first edition, so it has only been possible to make small changes. Fortunately these were mostly very minor but we should nonetheless like to draw the reader's attention to the more important alterations we have made.

The definition of $\mathscr{M}^*$ on page 214 has been changed to make the remarks on page 226 correct. The original proof of theorem 10.4.2 was right but we have had to modify it to square with the new definition of $\mathscr{M}^*$. Also, the cardinal calculations in lemmas 10.2.2 and 10.3.3 are only correct under conditions more restrictive than those originally specified. This has meant amending the hypotheses in various theorems (mostly in chapter 10) which follow from them.

We are very grateful to all those readers who have pointed out the mistakes in the first printing. We are particularly indebted to H. C. Doets and G. Fuhrken for their many penetrating observations and most of all we should like to thank Wilfrid Hodges for reading the book more carefully than we ever did and providing us with a rich source of criticisms and suggestions. The nature of this reprinting has prevented us from acting on all of the many suggestions received; and we alone are responsible for the errors that (we expect) remain. Readers are again encouraged to let us know of any they detect.

JOHN BELL
ALAN SLOMSON

Spring 1971

INTRODUCTION

"A book should have either intelligibility or correctness; to combine the two is impossible" – BERTRAND RUSSELL [1901].

The aim of this book is to provide an elementary exposition of some of the basic concepts of model theory. Model theory, which can be described briefly as the study of the relationship between formal languages and abstract structures, covers a very wide field and it is not possible to compress it into one volume. We have chosen as our theme the ultraproduct construction, and for the most part we have restricted our attention to those areas of the subject where this construction is of use.

We hope that this book will be of interest to undergraduates and practising mathematicians who would like to know about recent developments in the subject, but the intended audience is that of first year graduate students and our treatment has been chosen with them in mind. For example, in our experience very few undergraduates come into contact with any but the trivial parts of the theory of Boolean algebras and so our chapter on these is quite detailed. On the other hand most undergraduates will already have met propositional and predicate calculus and so our treatment of these, though self-contained, is fairly rapid.

The only previous knowledge we assume is some acquaintance with the rudiments of set theory. These prerequisites are described in ch. 0. There are also places where we use illustrations whose understanding will require some knowledge of topology or abstract algebra, but the general development of the book does not depend on these asides.

In ch. 1 we develop the notion of a Boolean algebra beginning with the idea of a partially ordered set, and we prove the results about Boolean algebras that we need later in the book. Ch. 2 is devoted to propositional

calculus and ch. 3 to predicate calculus. We give completeness proofs for the systems we introduce using the Boolean algebra methods of Rasiowa and Sikorski.

Model theory proper begins in ch. 4, where we introduce the basic ideas and prove some of the fundamental results, including the Löwenheim–Skolem theorems in all directions. In ch. 5 we introduce the ultraproduct construction which occupies most of our attention for the rest of the book. After proving Łoš's theorem, which is fundamental to everything else we do, we give some applications, beginning with a discussion of finite axiomatizability and ending with a proof of the compactness theorem. We interrupt our development of model theory in ch. 6 to give a detailed discussion of some of the set theoretic properties of the ultraproduct construction.

In ch. 7 we describe the various kinds of elementary classes. We obtain Keisler's theorem which characterizes elementary equivalence in terms of the ultrapower construction and we use it to provide algebraic descriptions of the elementary classes. Kochen's ultralimit theorem, which we prove in ch. 8, enables us to give similar characterizations which do not depend on the generalized continuum hypothesis (G.C.H.). Ch. 9 is devoted to completeness and model completeness. We use ultraproducts to derive Abraham Robinson's test for model completeness and then we give some applications to particular algebraic systems.

Ch. 10 is used to describe the work of Jónsson and of Morley and Vaught on homogeneous universal models, and in ch. 11 we use some results of Keisler to connect this up with the ultraproduct construction via the notion of a saturated structure.

We begin ch. 12 with a description of some classical results of Gödel and Skolem in the light of the ultraproduct construction. This leads in a natural way to more recent results on two cardinal theorems and cardinal like models.

The last two chapters are devoted to extensions of classical first order logic. In ch. 13 we describe some recent work on generalized quantifiers and in ch. 14 we discuss infinitary languages. Both chapters have as their major theme compactness results for these extended languages. Ch. 14 concludes with two applications of ultraproducts to questions about measurable cardinals.

The book contains a number of exercises, particularly in the early chapters. These are, we hope, all very easy, and are designed to test the reader's understanding of what has gone before, as well as to save us the chore of writing out long proofs of simple results.

The late Clifford Allen was once driven to make the following entry in his diary:

"More intellectual 'ticking off' from B.[ertrand] R.[ussell] at dinner because I used the word 'sentence' when I should have used 'phrase'. I'm dead sick of it."

GILBERT [1965], page 134.

Although there is a temptation for mathematical writers and particularly logicians to be excessively pedantic, we hope that we will not cause Allen's reaction in our readers. We have tried to avoid obscuring basically simple ideas under layers of precision. For example, we do not usually make the distinction between mention and use. It seems better to ignore this distinction rather than distract the reader by drawing his attention to it.

We are greatly indebted to all those logicians whose work we have tried to expound. We have attempted to give references for all results but we realize that, especially in the case of the better known results, there must be some omissions or misattributions for which we apologise in advance. Also it should be noted that in naming theorems we have been motivated more by the desire to make cross reference easy than by the desire to achieve historical accuracy.

We also have a number of personal debts to various colleagues and friends who have made useful suggestions. In particular we owe a great deal to Dr J. N. Crossley. He asked us to give a course of lectures on model theory in Oxford in 1965–66 and these lectures are the original source of this book. But for his continual encouragement and advice this book would never have appeared. Whether he is to be praised or blamed for this we leave to our readers to judge, but we at any rate are extremely grateful.

We should like to thank Jane Bridge and George Wilmers for their help with the proof reading. We are also grateful to our publishers, and especially their editor Mr W. P. Vlugt, for the care and skill with which the typescript has been turned into print.

We leave it to the reader to decide on which horn (if any!) of Russell's dilemma we have impaled ourselves. Although we have attempted to reduce mistakes, evasions and obscurities to a minimum we should be glad to have our attention drawn to any indiscretions the reader may discover in the text.

CHAPTER 0

PREREQUISITES

In this short chapter we list those topics, mostly taken from naive set theory, that we expect our readers to be already familiar with. We begin with two points of notation.

The book is divided into chapters and the chapters into sections. When we want to refer to a theorem, exercise etc. within the same chapter we give the number of the section in which it occurs, followed by its number in that section. When we refer to a result from another chapter we precede these numbers by the chapter number. Thus theorem 5.4.1 is the first result of §4 of ch. 5 and within ch. 5 is just referred to as theorem 4.1. We have given names to most of the theorems we need to refer to frequently and an index of named theorems is provided.

References to the bibliography are given thus: GÖDEL [1931]. The name of the author is followed by the year of original publication in square brackets. In cases where the same author has published more than one item that we need to refer to in the same year, the second item is distinguished by adding 'i' after the date, and so on.

We expect our readers to be familiar with the basic elements of naive set theory. The best account of this subject is ROTMAN and KNEEBONE [1966] and the contents of this work coincide almost exactly with what we assume. More advanced accounts include FRAENKEL [1966], SIERPINSKI [1965], KURATOWSKI and MOSTOWSKI [1967] and, in German, BACHMANN [1967]. Here we just mention some of the less elementary points.

We assume that the reader is familiar with the axiom of choice and its various equivalent formulations, in particular, that it is equivalent to Hausdorff's maximal principle, to Zorn's lemma, to the well-ordering principle and to the fact that all infinite cardinals are alephs. (In the absence of the axiom of choice, by an *infinite* set we mean a set into which the natural numbers can be mapped one–one.) For proofs of these equivalences consult RUBIN and RUBIN [1963].

We conceive of ordinals in such a way that each ordinal is the set of all smaller ordinals, and the finite ordinals as being identical with the natural

numbers. The order type of the natural numbers is ω and, by the previous remark, is the set of all natural numbers.

An ordinal is a *successor ordinal* if it is of the form $\xi+1$, otherwise it is a *limit ordinal*. Note that $\xi+1=\xi\cup\{\xi\}$, and that if ξ,ζ are ordinals, $\xi<\zeta$ is equivalent to $\xi\in\zeta$. We write $\xi<\zeta$ or $\xi\in\zeta$ according to which seems more natural in the context.

We regard cardinals as being identical with initial ordinals. So ω is also the cardinal of the natural numbers. This ambiguity will cause no confusion, but often, when it seems more natural, we use $\aleph_0$ for the cardinal of the natural numbers and N for the set of natural numbers.

If α is a cardinal, α^+ is the least cardinal greater than α. A cardinal is a *successor cardinal* if it is of the form α^+, otherwise it is a *limit cardinal*. Note that all (infinite) cardinals are limit ordinals.

An increasing sequence $\langle\gamma_\xi:\xi<\beta\rangle$ of ordinals is said to be *cofinal* in the ordinal α if for each ordinal $\delta<\alpha$, there is some $\xi<\beta$ such that $\delta\leqslant\gamma_\xi$. The *cofinality* of α, written $\mathrm{cf}(\alpha)$, is the least ordinal β such that there is an increasing sequence $\langle\gamma_\xi:\xi<\beta\rangle$ of ordinals each less than α, which is cofinal in α. The cofinality of α is always a cardinal and $\mathrm{cf}(\mathrm{cf}(\alpha))=\mathrm{cf}(\alpha)$.

An ordinal α is said to be *regular* if $\mathrm{cf}(\alpha)=\alpha$. It is *singular* if $\mathrm{cf}(\alpha)<\alpha$. Clearly there is no other possibility. It follows that if α is a cardinal, α is singular iff any set of cardinal α can be written as the union of fewer than α sets each of cardinal less than α. All successor cardinals are regular and hence all singular cardinals are limit cardinals. A regular limit cardinal is called an *inaccessible* cardinal. Thus $\aleph_0$ is an inaccessible cardinal. It is not possible to prove the existence of any other inaccessibles in ZF set theory. Note that some writers define inaccessibles so as to exclude $\aleph_0$.

If α and β are cardinals, α^β denotes the result of *cardinal* exponentiation and never the result of ordinal exponentiation. Thus 2^α is the cardinal of the set of all subsets of a set of cardinal α. If A and B are sets, A^B is the set of all functions defined on B with values in A. Since cardinals are sets, this notation can be ambiguous, but this should never lead to confusion. If ξ is an ordinal and A is a set, a *ξ-termed sequence* of elements of A is a member of the set A^ξ. Such a sequence f will often be written as

$$f=\langle f_\zeta:\zeta<\xi\rangle.$$

A cardinal α is said to be a *strong limit cardinal* if $\beta<\alpha$ implies $2^\beta<\alpha$. Regular strong limit cardinals are called *strongly inaccessible* cardinals. The *limit cardinal hypothesis*, abbreviated as L.C.H., is the hypothesis that

all singular cardinals are strong limit cardinals. The *generalized continuum hypothesis*, abbreviated as G.C.H., is the hypothesis that, for each infinite cardinal α,

$$2^{\alpha}=\alpha^{+}.$$

The G.C.H. is equivalent to the hypothesis, that for each ordinal ξ

$$2^{\aleph_\xi}=\aleph_{\xi+1}.$$

The *beths*, or *beth numbers*, $\beth_\xi$ where ξ is an ordinal, are defined recursively as follows:

$$\beth_0=\aleph_0,$$

if $\xi=\zeta+1$ is a successor ordinal

$$\beth_{\zeta+1}=2^{\beth_\zeta},$$

and if δ is a limit ordinal $\beth_\delta$ is the least cardinal greater than all the cardinals $\beth_\xi$ with $\xi<\delta$. A beth number is thus a cardinal which is equal to $\beth_\xi$ for some ξ. The G.C.H. implies that $\beth_\xi=\aleph_\xi$ for each ordinal ξ, and so the G.C.H. also implies that all (infinite) cardinals are beth numbers since we already know, as a consequence of the axiom of choice that all cardinals are alephs.

If α and β are cardinals then

$$\alpha^{<\beta}=\sum_{\delta<\beta}\alpha^{\delta},$$

where the sum is taken over all *cardinals* $\delta<\beta$.

If $f:A\to B$ is a function from a set A to a set B, $X\subseteq A$ and $Y\subseteq B$, then

$$f[X]=\{f(x):x\in X\} \quad \text{and} \quad f^{-1}[Y]=\{x:f(x)\in Y\}.$$

If X is a set we denote its cardinal by $\mathrm{card}(X)$, and its *power set* by $\mathscr{P}(X)$. Thus

$$\mathscr{P}(X)=\{A:A\subseteq X\}.$$

If α is a cardinal, $S_\alpha(X)$ is the set of all subsets of X of cardinal less than α and $\bar{S}_\alpha(X)$ is the set of all complements of the sets in $S_\alpha(X)$. Thus

$$Y\in\bar{S}_\alpha(X) \quad \text{iff} \quad Y\subseteq X \text{ and } \mathrm{card}(X-Y)<\alpha.$$

Finally notice that in common with most contemporary mathematical writers we use the very convenient abbreviation 'iff' for 'if and only if'. However 'off' means just off, whenever it occurs.

CHAPTER 1

BOOLEAN ALGEBRAS

In this book we shall be interested in Boolean algebras for two reasons. First, as we show in subsequent chapters, the structure of classical logic is essentially that of a Boolean algebra and so we can use theorems about Boolean algebras to deduce consequences about propositional and predicate logic. Second, the ultraproduct construction, which gets a great deal of attention in this book, makes use of ultrafilters in power set Boolean algebras and so we shall need to know something about them. In this chapter we explain what a Boolean algebra is and we prove the theorems about them that we shall need later on.

§1. Lattices

We begin with the notion of a *partially ordered set* which is an ordered pair $\langle X, \leqslant \rangle$ where X is a non-empty set and $\leqslant$ is a binary relation on X with the following properties:

(a). *Reflexivity:* for all $x \in X$, $x \leqslant x$,

(b). *Antisymmetry:* for all $x, y \in X$, if $x \leqslant y$ and $y \leqslant x$, then $x = y$,

(c). *Transitivity:* for all $x, y, z \in X$, if $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$.

A partially ordered set $\langle X, \leqslant \rangle$ is said to be *totally ordered* if, in addition, $\leqslant$ has the following property:

(d). *Dichotomy:* for all $x, y \in X$, either $x \leqslant y$ or $y \leqslant x$.

EXAMPLES

1.1. Let X be any set. We denote the *power set* of X, i.e. the set of all subsets of X, by $\mathscr{P}(X)$. If $\subseteq$ is the relation of set inclusion, then $\langle \mathscr{P}(X), \subseteq \rangle$ is a partially ordered set.

1.2. Let N be the set of natural numbers and let $\leqslant$ be the usual ordering of the natural numbers. Then $\langle N, \leqslant \rangle$ is a totally ordered set.

1.3. Again let N be the set of natural numbers. We define the relation $|$ on N by $m | n$ iff m divides n. Then $\langle N, | \rangle$ is a partially ordered set.

EXERCISES

1.4. In what circumstances is the partially ordered set $\langle \mathscr{P}(X), \subseteq \rangle$ of example 1.1 totally ordered?

1.5. Let $\leqslant$ be a binary relation on a non-empty set X. Define $<$ by $x<y$ iff $x \leqslant y$ and $x \neq y$. Find a set of conditions on $<$ necessary and sufficient for $\langle X, \leqslant \rangle$ to be a totally ordered set.

If $\langle X, \leqslant \rangle$ is a totally ordered set and $<$ is the relation defined as in exercise 1.5, then $<$ is called the *strict ordering* of X induced by $\leqslant$. Very often we shall think of totally ordered sets as consisting of sets equipped with a strict ordering. There is no essential difference between these two points of view but sometimes one is more convenient than the other.

For the rest of this chapter we will not mention the order relation $\leqslant$ explicitly unless we need to do so to avoid ambiguity. So we shall talk about 'the partially ordered set X' where the order relation on X is supposed to be $\leqslant$ unless we say otherwise. If $x \leqslant y$ we often use intuitive language and say 'y is greater than x'.

If X is a partially ordered set, a *chain* in X is a subset of X which is totally ordered by the order relation on X.

If A is a subset of the partially ordered set X, an *upper bound* for A in X is an element of X greater than all the elements of A. So $x \in X$ is an upper bound for A in X iff $a \leqslant x$ for all $a \in A$. Similarly $x \in X$ is a *lower bound* for A in X if for all $a \in A$, $x \leqslant a$.

$x \in X$ is called the *supremum* (or *least upper bound*) of A if x is an upper bound for A in X and is a lower bound for the set of all upper bounds for A in X. Thus $x \in X$ is the supremum of A iff for all $a \in A$, $a \leqslant x$ and given any other upper bound, say y, for A in X, $x \leqslant y$. Similarly $x \in X$ is said to be the *infimum* (or *greatest lower bound*) of A if it is a lower bound for A and greater than any other lower bound for A in X.

We denote the supremum of A, if it exists, by $\sup(A)$ and the infimum of A, if it exists, by $\inf(A)$. If $A = \{a_i : i \in I\}$ is an indexed collection of elements in X we often write $\bigvee_{i \in I} a_i$ for $\sup(A)$ and $\bigwedge_{i \in I} a_i$ for $\inf(A)$. If A is the two element set $\{x, y\}$, then we write $x \vee y$ for $\sup(A)$ and $x \wedge y$ for $\inf(A)$. $x \vee y$ is called the *join* of x and y, and $x \wedge y$ is called the *meet* of x and y.

A *lattice* is a partially ordered set in which each two element subset has both a supremum and an infimum. Thus in a lattice L for any $x, y \in L$ both $x \vee y$ and $x \wedge y$ exist.

EXAMPLE 1.6. The partially ordered set $\langle \mathscr{P}(X), \subseteq \rangle$ of example 1.1 is a

lattice for any set X. In this case for any $A,B\in\mathscr{P}(X)$, the supremum of $\{A,B\}$ is just $A\cup B$, the union of A and B, and its infimum is $A\cap B$, the intersection of A and B.

EXERCISES

1.7. Show that each totally ordered set is a lattice.

1.8. Prove that each finite non-empty subset of a lattice has both a supremum and an infimum.

Let $\mathfrak{m}$ be an infinite cardinal. An $\mathfrak{m}$-*complete* lattice is a lattice in which each set of elements of cardinal $\leqslant\mathfrak{m}$ has both an infimum and a supremum. A lattice is $<\mathfrak{m}$-*complete* if each set of elements of cardinal $<\mathfrak{m}$ has both an infimum and a supremum. A lattice is *complete* if it is $\mathfrak{m}$-complete for each infinite cardinal $\mathfrak{m}$. The previous exercise shows that each lattice is $<\aleph_0$-complete.

EXERCISES

1.9. Find an example of a partially ordered set that is not a lattice.

1.10. Prove that in any lattice L, for any $x,y,z\in L$ the following identities hold:

L_1. $\quad x\vee y=y\vee x \qquad x\wedge y=y\wedge x.$

L_2. $\quad x\vee(y\vee z)=(x\vee y)\vee z \qquad x\wedge(y\wedge z)=(x\wedge y)\wedge z.$

L_3. $\quad (x\vee y)\wedge y=y \qquad (x\wedge y)\vee y=y.$

1.11. If X and Y are finite subsets of a lattice L show that

$$\inf(X)\wedge\inf(Y)=\inf(X\cup Y).$$

Clearly, if a partially ordered set has an upper bound then it is unique. In particular this is true for a lattice and we call the unique upper bound of a lattice, if it exists, the *maximum* of the lattice and denote it by 1. Similarly if a lattice has a lower bound this lower bound is unique. We call it the *minimum* of the lattice and denote it by 0.

The maximum element of a lattice must be carefully distinguished from a *maximal* element of a lattice. An element x of a lattice is said to be *maximal* if there is no element of the lattice strictly greater than x. We remark here that *Zorn's lemma* states that if X is a partially ordered set in which each chain has an upper bound then X contains a maximal element. We shall be using this form of the axiom of choice later on. A partially ordered set does not necessarily have a unique maximal element; the next lemma tells us that the situation for lattices is different.

LEMMA 1.12. *If the lattice L contains a maximal element then this maximal element is unique.*

PROOF. Suppose that x and y are both maximal elements of L. Then $x \leqslant x \vee y$ and $y \leqslant x \vee y$ so that, since x and y are maximal, $x = x \vee y$ and $y = x \vee y$. Thus $x = y$ and the lemma follows.

A lattice L is said to be *complemented* if it has a maximum element 1, a minimum element 0, and for each $x \in L$ there is an element $y \in L$ satisfying:

L_4. $\quad x \vee y = 1 \quad$ and $\quad x \wedge y = 0$.

Such an element y is called a *complement* of x. In a complemented lattice each element does not necessarily have a unique complement. Consider, for example, the lattice whose underlying set is $X = \{0,a,b,c,1\}$ and whose order relation $\leqslant$ is given by:

$$\text{for all } x, y \in X, \quad x \leqslant y \quad \text{iff} \quad x = 0 \text{ or } y = 1 \text{ or } x = y.$$

Then, since in this lattice $a \vee b = b \vee c = c \vee a = 1$ and $a \wedge b = b \wedge c = c \wedge a = 0$, each one of the elements a,b,c is a complement of the other two.

The extra condition we impose on a complemented lattice to ensure the uniqueness of complements is distributivity. A lattice L is said to be *distributive* if for all $x,y,z \in L$ the following identities hold:

L_5. $\quad (x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z),$
$\quad\quad (x \wedge y) \vee z = (x \vee z) \wedge (y \vee z).$

EXERCISE 1.13. Show that in a complemented lattice L, for all $x \in L$, $x \wedge 1 = x$, $x \vee 1 = 1$, $x \wedge 0 = 0$, and $x \vee 0 = x$.

LEMMA 1.14. *In a distributive complemented lattice each element has a unique complement.*

PROOF. Suppose that L is a distributive complemented lattice and that y and z are both complements of the element x of L. So $x \vee y = 1$ and $x \wedge z = 0$. Therefore, using the results of exercise 1.13,

$$\begin{aligned} y &= y \vee 0 = y \vee (x \wedge z) \\ &= (y \vee x) \wedge (y \vee z), \quad \text{since } L \text{ is distributive}, \\ &= 1 \wedge (y \vee z) \\ &= y \vee z. \end{aligned}$$

Similarly $z = y \vee z$, and hence $y = z$. This completes the proof.

EXERCISE 1.15. Prove that if the first identity of L_5 holds in a lattice L for any elements x,y,z of the lattice, then so also does the second identity, and vice versa.

In a distributive complemented lattice the unique complement of an element x is denoted by x^*.

EXERCISE 1.16. Show that in a distributive complemented lattice, for each element x of the lattice $(x^*)^* = x$.

We are now ready for the chief definition of this chapter. A *Boolean algebra* is a complemented distributive lattice with at least two elements.*

We have chosen to introduce the notion of a Boolean algebra by building up its structure gradually starting with a partially ordered set. Alternatively we could have defined a Boolean algebra $\mathscr{B}$ to be an algebraic structure, say

$$\mathscr{B} = \langle B, \vee, \wedge, ^*, 0, 1 \rangle,$$

satisfying L_1–L_5. This explains why the term 'algebra' is used. If we had introduced Boolean algebras in this way, we could define the relation $\leqslant$ on B by $x \leqslant y$ iff $x \vee y = y$.

EXERCISE 1.17. Suppose $\mathscr{B} = \langle B, \vee, \wedge, ^*, 0, 1 \rangle$ is a structure satisfying L_1–L_5. Let $\leqslant$ be defined as above. Show that $\langle B, \leqslant \rangle$ is a complemented distributive lattice.

EXAMPLES

1.18. Let **2** be the set whose only elements are the cardinals 0,1 with the partial ordering $\leqslant$ defined by $0 \leqslant 0$, $0 \leqslant 1$, and $1 \leqslant 1$. Clearly **2** is a Boolean algebra. It is also a complete lattice.

1.19. The partially ordered set $\langle \mathscr{P}(X), \subseteq \rangle$ of example 1.1 is also a Boolean algebra. In this case the maximum element is the set X, and the minimum element is the empty set $\emptyset$, and complementation corresponds to set theoretic complementation. A Boolean algebra of the form $\langle \mathscr{P}(X), \subseteq \rangle$ is called a *power set Boolean algebra.*

1.20. Let X be a set. A subset A of X is said to be *cofinite* if $X - A$ is finite. Let $Z(X)$ be the set of all finite subsets and all cofinite subsets of X. Then $\langle Z(X), \subseteq \rangle$ is a Boolean algebra. If X is of cardinal $\aleph_0$ then $Z(X)$ also has cardinal $\aleph_0$. Since for any set Y, $\mathscr{P}(Y)$ is either finite or uncountable, $Z(X)$ cannot be isomorphic to a power set Boolean algebra.

* Thus we insist that $0 \neq 1$ in any Boolean algebra.

1.21. A subset of a topological space is said to be *clopen* if it is both open and closed. The family of clopen subsets of a topological space T is a Boolean algebra, called the *characteristic algebra* of T. Of course if T is a connected space, the characteristic algebra of T is isomorphic to the Boolean algebra **2** of example 1.18.

EXERCISE 1.22. Let N be the set of natural numbers. We define the relation $\sim$ on $\mathscr{P}(N)$ by

$$A \sim B \quad \text{iff} \quad (A-B)\cup(B-A) \text{ is finite}.$$

It is easily seen that $\sim$ is an equivalence relation on $\mathscr{P}(N)$. If $A \subseteq N$ let $A^{\sim}$ be the equivalence class to which A belongs under this relation and let $X=\{A^{\sim}: A \subseteq N\}$.

We define the relation $\leqslant$ on X by

$$A^{\sim} \leqslant B^{\sim} \quad \text{iff} \quad A-B \text{ is finite}.$$

Show that this is an unambiguous definition and that $\langle X, \leqslant \rangle$ is a Boolean algebra.

§ 2. Filters

A *filter* in a lattice L is a non-empty subset F of L which satisfies:

(a). for all $x, y \in F$, $x \wedge y \in F$,

(b). for all $x \in F$ and $y \in L$, if $x \leqslant y$ then $y \in F$.

An *ideal* in a lattice L is a non-empty subset I of L which satisfies:

(a). for all $x, y \in I$, $x \vee y \in I$,

(b). for all $x \in I$ and $y \in L$, if $y \leqslant x$ then $y \in I$.

It is clear that if a lattice L has a maximum element 1 then each filter in L contains it. Similarly the minimum of L, if it exists, is an element of each ideal in L.

EXAMPLES

2.1. The lattice L itself is both a filter and an ideal.

2.2. If 1 is the maximum of L and 0 the minimum of L then $\{1\}$ is a filter and $\{0\}$ is an ideal in L.

2.3. For each $x \in L$, the set $\{y: x \leqslant y\}$ is a filter, called the *principal filter* generated by x. Similarly $\{y: y \leqslant x\}$ is an ideal, the *principal ideal* generated by x.

EXERCISE 2.4. Show that in a finite lattice each filter and each ideal is principal, i.e. generated by some element of the lattice.

A filter F in a lattice L is said to be *proper* if F is a proper subset of L, that is if $F \neq L$. Clearly if L has a minimum 0, F is proper iff $0 \notin F$. Similarly an ideal I of a lattice is proper if $I \neq L$, and in a lattice with a maximum 1, I is proper iff $1 \notin I$. From now on by 'filter' and 'ideal' we will mean 'proper filter' and 'proper ideal' unless we say otherwise.

Duality. It can be seen that if in L_1–L_5 we interchange $\wedge$ and $\vee$, and 0 and 1, the identities are transformed into new identities which also hold in all Boolean algebras. In fact, as we have set out L_1–L_5 the two identities of each pair become interchanged under this transformation. The statement obtained from making this transformation is called the *dual* of the original statement. It follows that the dual of any true statement about Boolean algebras expressed in terms of $\wedge$, $\vee$, 0 and 1 is also a true statement about Boolean algebras. When the statement also contains mention of ideals and filters we obtain its dual by also interchanging 'ideal' and 'filter', and '$\leqslant$' and '$\geqslant$'. (We write $y \geqslant x$ for $x \leqslant y$.) As an example consider the following lemma.

LEMMA 2.5. *In a Boolean algebra* $x \wedge y^* = 0$ *iff* $x \leqslant y$.

PROOF. Suppose $x \wedge y^* = 0$. Then $x = x \wedge 1 = x \wedge (y \vee y^*) = (x \wedge y) \vee (x \wedge y^*) = (x \wedge y) \vee 0 = x \wedge y \leqslant y$.

Conversely if $x \leqslant y$, then $x \wedge y = x$ and so $x \wedge y^* = (x \wedge y) \wedge y^* = x \wedge (y \wedge y^*) = x \wedge 0 = 0$.

Now by duality we can immediately deduce that *in a Boolean algebra* $x \vee y^* = 1$ *iff* $y \leqslant x$.

In future, when we have proved a theorem about Boolean algebras, we will assume that we have also established its dual.

A set A of elements of a Boolean algebra is said to have the *finite intersection property* (henceforth abbreviated as *'fip'*) if the infimum of any finite subset of A is not equal to 0. The importance of this property is that it is precisely those subsets of a Boolean algebra that have the fip that can be extended to filters, as we now see.

EXERCISES

2.6. Show that if A has the fip, then for any element x of the Boolean algebra either $A \cup \{x\}$ or $A \cup \{x^*\}$ has the fip.

2.7. Let $\{A_i : i \in I\}$ be a chain of subsets of a Boolean algebra (i.e. a collection of subsets totally ordered by set inclusion). Show that if each A_i has the fip then so has the union.

If A is a set of elements of a Boolean algebra B, we let A^0 be the set of those elements of B greater than some element of A. That is

$$A^0 = \{x \in B : \text{for some } a \in A,\ a \leqslant x\}.$$

We let A^c be the set of the infimums of all the finite subsets of A. That is

$$A^c = \{\inf(X) : X \in S_\omega(A)\}.$$

(Recall that $S_\omega(A)$ is the set of all finite subsets of A.)

A *base* for a filter F is a set A such that $A^0 = F$. A *sub-base* for a filter is a set A such that A^c is a base for F. If A is a sub-base for F then $F = (A^c)^0$, and we say that A *generates* F. Note that $A \subseteq A^c \subseteq (A^c)^0$.

LEMMA 2.8. *For any subset A of a Boolean algebra, the set $(A^c)^0$ is a (not necessarily proper) filter. Any filter containing A contains $(A^c)^0$. $(A^c)^0$ is proper iff A has the fip.*

PROOF. Note first that $x \in (A^c)^0$ iff for some $X \in S_\omega(A)$, $\inf(X) \leqslant x$.

Suppose $x, y \in (A^c)^0$. Then for some $X, Y \in S_\omega(A)$, $\inf(X) \leqslant x$ and $\inf(Y) \leqslant y$. $X \cup Y \in S_\omega(A)$ and by exercise 1.11, $\inf(X \cup Y) = \inf(X) \wedge \inf(Y)$, but $\inf(X) \wedge \inf(Y) \leqslant x \wedge y$ and so $x \wedge y \in (A^c)^0$. Trivially if $x \in (A^c)^0$ and $x \leqslant y$ then $y \in (A^c)^0$. So $(A^c)^0$ is a filter.

Suppose F is an filter and $A \subseteq F$. Let $x \in (A^c)^0$. Then for some $X \in S_\omega(A)$, $\inf(X) \leqslant x$. Since $X \subseteq F$ and F is a filter $\inf(X) \in F$. Hence, again using the fact that F is a filter, $x \in F$. Thus $(A^c)^0 \subseteq F$.

If A does not have the fip for some $X \in S_\omega(A)$, $\inf(X) = 0$. Therefore $0 \in (A^c)^0$ and so $(A^c)^0$ is not proper. Now suppose $(A^c)^0$ is not proper. Then for some $X \in S_\omega(A)$, $\inf(X) \leqslant 0$ and so $\inf(X) = 0$. Thus A does not have the fip.

This completes the proof.

From this lemma it follows that a subset A of a Boolean algebra can be extended to a (proper) filter iff it has the fip.

§ 3. Ultrafilters

Filters in a Boolean algebra B are subsets of B with certain properties. As subsets they can be partially ordered by set inclusion. Filters which are maximal with respect to this ordering are called *ultrafilters*. Thus an ultrafilter is a filter F which has no proper extension which are also (proper) filters. Ultrafilters can also be characterized in another way, as shown by the next lemma.

LEMMA 3.1. *If F is a filter in a Boolean algebra B, F is an ultrafilter iff for each $x\in B$ either $x\in F$ or $x^*\in F$, but not both.*

PROOF. Note first that if for some $x\in B$ both x and x^* are in F, then $0=x\wedge x^*\in F$ and so F is not proper.

Now suppose that for each $x\in B$ either x or x^* is in F. Let G be a filter properly containing F. Then there is some x in $G-F$. Since $x\notin F$, $x^*\in F\subseteq G$. This shows that G is not proper. Hence F is a maximal proper filter, i.e. an ultrafilter.

Conversely suppose that F is an ultrafilter and $x\notin F$. Let G be the filter generated by $F\cup\{x\}$. Since F is an ultrafilter G is not proper. Therefore $F\cup\{x\}$ does not have the fip and so for some $X\in S_\omega(F)$, $\inf(X)\wedge x=0$. Hence $\inf(X)\leqslant x^*$. $\inf(X)$ is in F and therefore we can conclude that $x^*\in F$, thus completing the proof.

EXAMPLE 3.2. Let $\langle\mathscr{P}(X),\subseteq\rangle$ be the Boolean algebra of example 1.19. Then for each $x\in X$, the set $\{A\in\mathscr{P}(X):x\in A\}$ is an ultrafilter. We call this ultrafilter the *principal ultrafilter generated by x*. An ultrafilter is *principal* if it is generated by some element of the Boolean algebra.

EXERCISE 3.3. If X is any set, show that an ultrafilter in $\langle\mathscr{P}(X),\subseteq\rangle$ is principal iff it contains a finite set. Hence deduce that if X is finite all ultrafilters in $\langle\mathscr{P}(X),\subseteq\rangle$ are principal. (Hint: see ch. 6, §1 for a proof of this.)

We now use the axiom of choice (in the form of Zorn's lemma) to establish the existence of an extremely rich class of ultrafilters in an arbitrary Boolean algebra.

THEOREM 3.4. THE ULTRAFILTER THEOREM

Every filter in a Boolean algebra can be extended to an ultrafilter.

PROOF. Let F be a filter in a Boolean algebra B. Let $\mathscr{F}$ be the set of all filters in B which contain F. Since $F\in\mathscr{F}$, $\mathscr{F}$ is not empty.

$\mathscr{F}$ can be partially ordered by set inclusion. We will show that with respect to this ordering chains in $\mathscr{F}$ have upper bounds.

Let $\mathscr{D}=\{D_i:i\in I\}$ be a chain in $\mathscr{F}$, and let $D=\bigcup_{i\in I}D_i$. If $x,y\in D$, then for some $i,j\in I$, $x\in D_i$ and $y\in D_j$. Since $\mathscr{D}$ is a chain either $D_i\subseteq D_j$ or $D_j\subseteq D_i$, say $D_i\subseteq D_j$. Then $x,y\in D_j$ and so because D_j is a filter $x\wedge y\in D_j\subseteq D$. If $z\in B$ and $x\leqslant z$, then $z\in D_j\subseteq D$. Since $0\notin D_i$ for any $i\in I$, it follows that $0\notin D$. These remarks show that D is a filter. $F\subseteq D$ and so $D\in\mathscr{F}$. D is an upper bound for $\mathscr{D}$ in $\mathscr{F}$.

We can therefore deduce from Zorn's lemma that $\mathscr{F}$ contains a maximal element G. G is the required ultrafilter which is an extension of F.

COROLLARY 3.5. *Each set of elements of a Boolean algebra having the fip can be extended to an ultrafilter.*

PROOF. This is an immediate consequence of lemma 2.8 and the ultrafilter theorem.

COROLLARY 3.6. *Each non-zero element of a Boolean algebra is contained in some ultrafilter.*

PROOF. If $x \neq 0$, $\{x\}$ has the fip and so can be extended to an ultrafilter.

COROLLARY 3.7. *If x, y are distinct elements of a Boolean algebra there is an ultrafilter containing one but not the other.*

PROOF. If $x \neq y$ then either not $x \leqslant y$ or not $y \leqslant x$. Say not $x \leqslant y$. Hence, by lemma 2.5, $x \wedge y^* \neq 0$. Therefore $\{x, y^*\}$ has the fip and so can be extended to an ultrafilter F in B. Since $y^* \in F$, $y \notin F$.

We use corollary 3.5 a great deal subsequently when we want to extend a subset A of a Boolean algebra to an ultrafilter. The corollary shows that to do this we need only prove that A has the fip. To do this we must first show that $0 \notin A$. It will not in general then be sufficient to show that the meet of any two (rather than of any finite number of) elements of A is not 0. But if we can show that the meet of any two elements of A is in A then we will have shown that A has the fip, for it will immediately follow that the meet of any finite number of elements of A is in A and hence is not 0. This is what we shall in fact do on most occasions when we make use of this corollary. An important example is given below.

LEMMA 3.8. *If X is an infinite set then the Boolean algebra $\langle \mathscr{P}(X), \subseteq \rangle$ contains a non-principal ultrafilter.*

PROOF. The set $\bar{S}_\omega(X)$ of all cofinite subsets of X has the finite intersection property. Certainly all sets in $\bar{S}_\omega(X)$ are non-empty. If $X_1, X_2 \in \bar{S}_\omega(X)$, then

$$X - (X_1 \cap X_2) = (X - X_1) \cup (X - X_2)$$

is finite. Thus $X_1 \cap X_2 \in \bar{S}_\omega(X)$. By the remarks above this proves that $\bar{S}_\omega(X)$ has the finite intersection property. It therefore follows from corollary 3.5 to the ultrafilter theorem that $\bar{S}_\omega(X)$ can be extended to an

ultrafilter F. This ultrafilter contains no finite sets and therefore, by exercise 3.3, must be non-principal.

EXERCISE 3.9. Construct a subset A of a Boolean algebra having the following properties: $0 \notin A$, the meet of any two elements of A is not 0, but A does not have the fip.

§4. Homomorphisms and quotient algebras

If B is a Boolean algebra, a *subalgebra* of B is a non-empty subset of B which is closed under the operations of meet, join and complementation. Suppose B' is a subalgebra of B. Then there is some $x \in B'$. Since B' is closed under complementation, $x^* \in B'$. Therefore, because B' is closed under meets and joins, $1 = x \vee x^* \in B'$, and $0 = x \wedge x^* \in B'$. Thus a subalgebra B' of B contains both 0 and 1. It can easily be seen that the subset $\{0,1\}$ of B is a subalgebra of B. We have already seen that $\{0\}$ is an ideal and $\{1\}$ is a filter. This observation is generalized by the following exercise.

EXERCISE 4.1. Let B be a Boolean algebra and let F be a filter in B. Let $I = \{x^* : x \in F\}$. Show that I is an ideal, $B' = F \cup I$ is a subalgebra of B and that F is an ultrafilter in B'.

If B_1, B_2 are two Boolean algebras a *homomorphism* of B_1 into B_2 is a map f of B_1 into B_2 which preserves the algebraic operations, i.e. such that for all $x, y \in B_1$:

(a). $f(x \wedge y) = f(x) \wedge f(y)$,

(b). $f(x \vee y) = f(x) \vee f(y)$,

and

(c). $f(x^*) = f(x)^*$.

EXERCISE 4.2. If f is a homomorphism of B_1 into B_2 prove that

(a). f maps the maximum of B_1 onto the maximum of B_2, and the minimum of B_1 onto the minimum of B_2.

(b). $x \leqslant y$ implies $f(x) \leqslant f(y)$, for any $x, y \in B_1$

(c). $f[B_1]$ is a subalgebra of B_2, where $f[B_1]$ is the image of B_1 under f, that is, the set of all those elements of B_2 onto which some element of B_1 is mapped by f.

If f is a one–one map we say that f is an *isomorphism* between B_1 and $f[B_1]$.

If f is also an onto map, i.e. if $f[B_1]=B_2$, then f is an isomorphism between B_1 and B_2. If there is an isomorphism between B_1 and B_2 we say that B_1 and B_2 are *isomorphic* and we write $B_1 \simeq B_2$.

Let F be a filter in a Boolean algebra B. We define the relation $\sim_F$ on B by

$$x \sim_F y \quad \text{iff for some} \quad f \in F,\ x \wedge f = y \wedge f\ .$$

LEMMA 4.3. *$\sim_F$ is an equivalence relation on B.*

PROOF. Since F is a filter, F is not empty and so $\sim_F$ is trivially reflexive. Clearly $\sim_F$ is symmetric, so it remains only to show that $\sim_F$ is a transitive relation.

Suppose $x \sim_F y$ and $y \sim_F z$. Then for some $f,g \in F$, $x \wedge f = y \wedge f$ and $y \wedge g = z \wedge g$. Since F is a filter $f \wedge g \in F$ and using the fact that $\wedge$ is an associative and commutative operation $x \wedge (f \wedge g) = (y \wedge f) \wedge g = (y \wedge g) \wedge f = z \wedge (f \wedge g)$. Hence $x \sim_F z$ and our proof is completed.

LEMMA 4.4. *$\sim_F$ is a congruence relation on B. That is if $x \sim_F x'$ and $y \sim_F y'$ then*

$$x \wedge y \sim_F x' \wedge y',$$
$$x \vee y \sim_F x' \vee y',$$

and

$$x^* \sim_F x'^*.$$

PROOF. Suppose $x \wedge f = x' \wedge f$, and $y \wedge g = y' \wedge g$, with $f, g \in F$. It is easily checked that it follows from these relations that

$$(x \wedge y) \wedge (f \wedge g) = (x' \wedge y') \wedge (f \wedge g)$$

and

$$(x \vee y) \wedge (f \wedge g) = (x' \vee y') \wedge (f \wedge g),$$

whence, since $f \wedge g \in F$,

$$x \wedge y \sim_F x' \wedge y' \quad \text{and} \quad x \vee y \sim_F x' \vee y'.$$

Also we have

$$f = f \wedge (x \vee x^*) = (f \wedge x) \vee (f \wedge x^*) = (f \wedge x') \vee (f \wedge x^*),$$

and so

$$\begin{aligned} x'^* \wedge f &= x'^* \wedge [(f \wedge x') \vee (f \wedge x^*)] \\ &= [x'^* \wedge (f \wedge x')] \vee [(x'^* \wedge x^*) \wedge f] \\ &= (x'^* \wedge x^*) \wedge f\ . \end{aligned}$$

Similarly $x^* \wedge f = (x'^* \wedge x^*) \wedge f$ and therefore $x^* \sim_F x'^*$.

This completes the proof.

EXERCISE 4.5. Let $x \odot y$ be an abbreviation for $(x \vee y^*) \wedge (x^* \vee y)$. Show that

(a). $x \sim_F y$ iff $x \odot y \in F$,

and

(b). $x \odot y = 1$ iff $x = y$.

We let $|x|$ be the equivalence class to which the element x of B belongs under the equivalence relation $\sim_F$. Lemma 4.4 shows that $B/F = \{|x| : x \in B\}$ can be given the structure of a Boolean algebra if we define for each $x, y \in B$

$$|x| \wedge |y| = |x \wedge y|,$$
$$|x| \vee |y| = |x \vee y|,$$
$$|x|^* = |x^*|.$$

The set B/F with this structure is called the *quotient algebra of B modulo F*. The map $h: B \to B/F$ which sends x onto $|x|$ is clearly a homomorphism. It is called the *canonical homomorphism* of B onto B/F.

EXERCISE 4.6. Show that $|x| = 1$ in B/F iff $x \in F$.

We have thus seen that associated with each filter in B there is a homomorphism. The next lemma shows that the converse is also true.

LEMMA 4.7. *Let $f: B_1 \to B_2$ be a homomorphism between two Boolean algebras. The set $F = \{x \in B_1 : f(x) = 1\}$ is a filter and $f[B_1]$ is isomorphic to B_1/F.*

PROOF. The verification that F is a filter is straightforward and is left to the reader.

To prove the second part we first note that, for any $x, y \in B_1$,

$\|x\| = \|y\|$ iff	$x \sim_F y$	
iff	$x \odot y \in F$,	by exercise 4.5
iff	$f(x \odot y) = 1$,	by the definition of F,
iff	$f(x) \odot f(y) = 1$,	since f is a homomorphism
iff	$f(x) = f(y)$,	by exercise 4.5

where $|x|$ is the image of x under the canonical homomorphism $h: B_1 \to B_1/F$. Therefore the map g from B_1/F onto $f[B_1]$ defined by

$$g(|x|) = f(x)$$

is well-defined and is one-one.

It is easy to see that g is a homomorphism of B_1/F onto $f[B_1]$. Therefore $B_1/F \simeq f[B_1]$.

We call the filter F the *shell* of the homomorphism f. The dual notion to that of a shell is that of the *kernel* of a homomorphism f, which is defined to be the set $\{x: f(x)=0\}$. It can be seen that the kernel of a homomorphism is an ideal. Lemma 4.7 is usually stated in its dual form, that is, in terms of ideals. In this case the term 'kernel' is well known. The nomenclature 'shell' was suggested to us by R. J. Baxter and arises from the obvious botanical analogy.

LEMMA 4.8. *A homomorphism is one–one iff its shell is* $\{1\}$.

PROOF. Let $f: B_1 \to B_2$ be a homomorphism between two Boolean algebras with shell $\{1\}$. Suppose $f(x)=f(y)$. Then $f(x \vee y^*)=f(x) \vee f(y^*)= =f(x) \vee f(y)^*=f(x) \vee f(x)^*=1$. Similarly $f(x^* \vee y)=1$ and so $f(x \odot y)=1$. It follows that $x \odot y=1$ and so, by exercise 4.5, $x=y$. This shows that f is one–one. The converse is trivial.

A filter F in a Boolean algebra is said to be *prime* if whenever $x \vee y \in F$ then either $x \in F$ or $y \in F$.

THEOREM 4.9. *The following conditions on a filter F in a Boolean algebra B are equivalent:*

(a). $B/F \simeq \mathbf{2}$ *(where* **2** *is defined as in example* 1.18*)*

(b). *F is an ultrafilter*

(c). *F is prime*

(d). *for any $x \in B$, either x or x^* is in F.*

PROOF. (b)$\Leftrightarrow$(d). This is just lemma 3.1.

(d)$\Rightarrow$(a). Let $|x|$ be the image of x under the canonical homomorphism of B onto B/F and suppose $|x| \neq 1$. Then by exercise 4.6, $x \notin F$, hence assuming (d) $x^* \in F$. So $|x|^*=|x^*|=1$. Thus $|x|=|x|^{**}=0$. This shows that

$$B/F=\{0,1\} \simeq \mathbf{2}.$$

(a)$\Rightarrow$(d). If $x \notin F$ then $|x| \neq 1$ and so, assuming (a), $|x|=0$ whence $|x^*|=1$ and therefore $x^* \in F$.

(b)$\Rightarrow$(c). Suppose $x \vee y \in F$, but $x \notin F$. Let G be the filter generated by $F \cup \{x\}$. Since, assuming (b), F is maximal, G is not proper and hence for some $z \in F$, $z \wedge x=0$. Since z, $x \vee y \in F$, $z \wedge (x \vee y)=(z \wedge x) \vee (z \wedge y)= =0 \vee (z \wedge y)=z \wedge y \in F$. But $z \wedge y \leqslant y$ and since F is a filter, $y \in F$.

(c)$\Rightarrow$(d). For any $x \in B$, $x \vee x^*=1 \in F$ and so, assuming (c) either $x \in F$ or $x^* \in F$.

We conclude this section with a result that we need later to prove the completeness of predicate calculus.

Let $\{A_n : n<\omega\}$ be a *countable* collection of subsets of a Boolean algebra each having an infimum. Say

$$\text{for } n<\omega \qquad a_n=\inf(A_n). \tag{Π}$$

An ultrafilter F in B is said to *preserve* the meets [Π] iff

$$\text{for } n<\omega \qquad h(a_n)=\inf\{h(a): a\in A_n\},$$

where h is the canonical homomorphism of B onto $B/F\simeq \mathbf{2}$.

THEOREM 4.10. TARSKI'S LEMMA

If $x\neq 0$ there is an ultrafilter F containing x which preserves the meets [Π].

PROOF. We define by recursion a sequence $\{b_n : n<\omega\}$ such that for each $n<\omega$, $b_n\in A_n$ and $\{x, a_0\vee b_0^*, \ldots, a_n\vee b_n^*\}$ has the fip.

Suppose $m<\omega$, and for $n<m$ we have found b_n to satisfy these conditions. If $m=0$, let $y=x$, and if $m>0$ let

$$y=x\wedge(a_0\vee b_0^*)\wedge\cdots\wedge(a_{m-1}\vee b_{m-1}^*).$$

In either case $y\neq 0$.

Suppose that for each $b\in A_m$, $y\wedge(a_m\vee b^*)=0$. Since

$$y\wedge(a_m\vee b^*)=(y\wedge a_m)\vee(y\wedge b^*)$$

it follows that $y\wedge a_m=0$ and $y\wedge b^*=0$ for each $b\in A_m$. Therefore $y\leqslant b$ for each $b\in A_m$ and so $y\leqslant a_m=\inf(A_m)$. Hence $y=y\wedge a_m=0$, contradicting our hypothesis. Therefore there is some $b_m\in A_m$ with

$$y\wedge(a_m\vee b_m^*)\neq 0.$$

Hence for $n\leqslant m$, we can find b_n so as to satisfy the desired conditions. Therefore we can find such a b_n for all $n<\omega$. It follows from these conditions that the set

$$\{x, a_0\vee b_0^*, \ldots, a_n\vee b_n^*, \ldots\}$$

has the fip and so can be extended to an ultrafilter F in B. We show that F preserves the meets [Π].

Let h be the canonical homomorphism of B onto B/F. Since $a_n\vee b_n^*\in F$, $h(a_n)\vee h(b_n)^*=h(a_n\vee b_n^*)=1$ and so $h(a_n)\geqslant h(b_n)$, whence

$$\inf\{h(b): b\in A_n\}\leqslant h(a_n). \tag{1}$$

Since $a_n=\inf(A_n)$, $a_n\leqslant b$ for each $b\in A_n$ and so $h(a_n)\leqslant h(b)$ for each $b\in A_n$. Therefore

$$h(a_n)\leqslant\inf\{h(b): b\in A_n\}. \tag{2}$$

We can conclude from (1) and (2) that

$$h(a_n)=\inf\{h(b): b\in A_n\}$$

and so F preserves the meets [Π].

EXERCISE 4.11. We use the notation of exercise 1.22. Let B be the Boolean algebra $\langle \mathscr{P}(N), \subseteq \rangle$ and let F be the filter of cofinite subsets of N. Show that

$$\langle X, \leqslant \rangle \simeq B/F .$$

§ 5. Filters in topology

In this section we discuss the role of filters in topology and in the next we prove Stone's representation theorem for Boolean algebras. Both these sections presuppose some knowledge of the elements of point set topology. However nothing later in the book depends on them and they can be omitted by the reader if he chooses.

If F is a filter in the power set Boolean algebra $\langle \mathscr{P}(X), \subseteq \rangle$ we say that F is a *filter on* X. Similarly an *ultrafilter on* X is an ultrafilter in $\langle \mathscr{P}(X), \subseteq \rangle$.

Let X be a topological space. If A is a subspace of X we denote its closure by $\bar{A}$. For each point $x\in X$ the family N_x of neighbourhoods of x is a filter on X, the *neighbourhood filter* of x.

LEMMA 5.1. *Let F be a filter on X. The following conditions are equivalent:*

(a). $x\in \bigcap\{\bar{A}: A\in F\}$.

(b). *for all $A\in F$ and $U\in N_x$, $A\cap U\neq\emptyset$.*

(c). *there is a proper filter containing $F\cup N_x$.*

PROOF. Clearly (b) and (c) are equivalent. If x is in the closure of each $A\in F$ then every neighbourhood of x meets each A in F and so (a) implies (b). The converse is similar.

A point x which satisfies the conditions in lemma 5.1 is said to be *adherent* to the filter F, or an *adherent point* of F. A filter F on X is said to *converge* to a point x if $N_x\subseteq F$. Condition (b) of lemma 5.1 shows that a point is adherent to every filter convergent to it. For ultrafilters these concepts are, in fact, equivalent.

LEMMA 5.2. *An ultrafilter converges to precisely its adherent points.*

PROOF. We have already remarked that an ultrafilter converges only to its adherent points. Now suppose x is an adherent point of an ultrafilter F.

Then by (c) of lemma 5.1, $F \cup N_x$ is contained in some filter G. Since F is an ultrafilter $F=G$ and so $N_x \subseteq F$, that is, F converges to x.

We are now able to characterize several important topological notions in terms of the convergence of filters.

LEMMA 5.3. *A space X is Hausdorff iff each filter on X converges to at most one point, or, equivalently, iff each ultrafilter has at most one adherent point.*

PROOF. X is Hausdorff iff for each distinct pair of elements $x, y \in X$, there is some $U \in N_x$ and $V \in N_y$ with $U \cap V = \emptyset$. Thus no filter contains both N_x and N_y and so each filter converges to at most one point.

LEMMA 5.4. *A space X is compact iff each filter on X has at least one adherent point.*

PROOF. Let F be a filter on a compact space X. The family $\bar{F} = \{\bar{A} : A \in F\}$ has the fip and therefore, since X is compact $\bigcap \bar{F}$ is not empty. Each $x \in \bigcap \bar{F}$ is an adherent point of F.

For the converse, suppose that F is a family of closed subsets of a space X which has the fip and suppose that each filter on X has at least one adherent point. F can be extended to a filter G. Since G has an adherent point, $\emptyset \neq \bigcap \{\bar{A} : A \in G\} \subseteq \bigcap \{\bar{A} : A \in F\} = \bigcap F$. Thus X is compact.

COROLLARY 5.5. *A space X is compact iff each ultrafilter on X converges to at least one point.*

PROOF. Assume X is compact. Then each filter and so each ultrafilter has at least one adherent point. Hence each ultrafilter converges to at least one point.

Conversely suppose that each ultrafilter converges to some point and let F be a filter on X. F can be extended to an ultrafilter G. G converges to some point x which is an adherent point of G and hence of F. Thus every filter has an adherent point and therefore X is compact.

§ 6. The Stone representation theorem

We know that not every Boolean algebra is isomorphic to a power set Boolean algebra. For example, as we noted earlier in example 1.20, the algebra of finite and cofinite subsets of a countable set is itself countable and so cannot be isomorphic to $\mathscr{P}(X)$ for any set X. However all is not lost; Stone's theorem, which we prove below, asserts that every Boolean algebra is isomorphic to a *subalgebra* of a power set Boolean algebra.

Let B be a Boolean algebra and let $S(B)$ be the set of all ultrafilters in B. If $X \subseteq S(B)$ we denote its complement in $S(B)$, $S(B)-X$, by cX.

THEOREM 6.1. THE STONE REPRESENTATION THEOREM
B is isomorphic to a subset of $\mathscr{P}S(B)$ considered as a power set Boolean algebra.

PROOF. We define the map $u: B \to \mathscr{P}S(B)$ as follows: for each $x \in B$, $u(x)$ is the set of all ultrafilters in B which contain x. We show that u is an isomorphism between B and $u[B]$.

Suppose $x, y \in B$ and $x \neq y$. By corollary 3.7 there is an ultrafilter containing one of these elements but not the other. Therefore $u(x) \neq u(y)$ and so u is one-one.

For any $x, y \in B$ and any filter F in B, $x \wedge y \in F$ iff $x \in F$ and $y \in F$; it follows that $u(x \wedge y) = u(x) \cap u(y)$. For any ultrafilter U, $x \in U$ iff $x^* \notin U$, and therefore $u(x^*) = cu(x)$.

These remarks show that u is an isomorphism as claimed.

We can go further than this last theorem. To do so we need first some topological preliminaries. A *Boolean space* is a compact Hausdorff space with a base of open and closed sets. Recall that we call a set which is both open and closed a *clopen* set. Thus a Boolean space is a compact Hausdorff space with a base of clopen sets. The elements of this base, being closed subsets of a compact space, are themselves compact.

A subspace of a topological space is *connected* if it is not the union of two non-empty disjoint open sets. A *component* of a topological space is a maximal connected subspace. The topological space X is said to be *totally disconnected* if each of the components of X contains just one point.

LEMMA 6.2. *A Boolean space is totally disconnected.*

PROOF. Let X be a Boolean space and suppose that A is a subspace of X containing two distinct points x and y. We show that A is not connected.

Since X is a Hausdorff space with a base of clopen sets, there is a clopen set U such that $x \in U$ and $y \notin U$. Then

$$A = (A \cap U) \cup (A - U)$$

and hence is not connected.

EXAMPLE 6.3. The set of all ordinals up to and including a given ordinal α, i.e. the set $\{\xi : \xi \leqslant \alpha\}$, is a Boolean space under the order topology (i.e. the

topology which has a sub-base consisting of the open ended intervals $\{\xi:\xi<\beta\}$ and $\{\xi:\beta<\xi\leqslant\alpha\}$).

Further examples will be given below.

We have already remarked that given a topological space X, the clopen subsets of X form a Boolean algebra with respect to the partial ordering of set inclusion. We call this Boolean algebra the *characteristic algebra* of X and denote it by $\boldsymbol{C}(X)$. $\boldsymbol{C}(X)$ is a subalgebra of the Boolean algebra $\langle\mathscr{P}(X),\subseteq\rangle$. In addition if X is a Boolean space then $\boldsymbol{C}(X)$ forms a base for the topology of X. The next results shows that in this case $\boldsymbol{C}(X)$ is uniquely characterized by these two properties.

LEMMA 6.4. *Let X be a Boolean space and let $\mathscr{A}$ be a subalgebra of $\langle\mathscr{P}(X),\subseteq\rangle$ which is a base for the topology of X. Then $\mathscr{A}=\boldsymbol{C}(X)$.*

PROOF. The elements of $\mathscr{A}$ are open subsets of X since they form a base for the topology. Since $\mathscr{A}$ is a subalgebra of $\langle\mathscr{P}(X),\subseteq\rangle$ it is closed under the formation of complements and hence the sets in $\mathscr{A}$ are also closed. Thus $\mathscr{A}\subseteq\boldsymbol{C}(X)$.

Now suppose $Y\in\boldsymbol{C}(X)$. Y is open and $\mathscr{A}$ is a base for the topology of X. Hence for each $y\in Y$ there is some $A_y\in\mathscr{A}$ with

$$y\in A_y\subseteq Y.$$

$\{A_y:y\in Y\}$ forms an open cover of Y. Y is a closed subset of a compact space and is therefore compact. It follows that this open cover has a finite subcover, say $\{A_{y_i}:1\leqslant i\leqslant n\}$. Thus

$$Y=A_{y_1}\cup\cdots\cup A_{y_n}.$$

But $\mathscr{A}$, being a subalgebra, is closed under finite unions. Therefore $Y\in\mathscr{A}$, and our proof that $\mathscr{A}=\boldsymbol{C}(X)$ is completed.

Now suppose B is a Boolean algebra. The set $S(B)$ of all ultrafilters in B can be given a topology in the following way. We take as a base for the topology the collection

$$u[B]=\{u(x):x\in B\}$$

of subsets of $S(B)$, where as before

$$u(x)=\{U\in S(B):x\in U\}.$$

Not surprisingly we now have

THEOREM 6.5. *The Stone space $S(B)$ of a Boolean algebra B is a Boolean space and the map u is an isomorphism of B onto the characteristic algebra $\boldsymbol{C}(S(B))$ of $S(B)$.*

PROOF. $S(B)$ is a Hausdorff space. For suppose $U,V\in S(B)$ and $U\neq V$. Then for some $x\in B$, $x\in U$ but $x\notin V$ and so $x^*\in V$. Then $U\in u(x)$ and $V\in u(x^*)$ and $u(x)$, $u(x^*)$ are disjoint open subsets of $S(B)$.

$S(B)$ is also a compact space. To show this it is sufficient to show that every cover of $S(B)$ by members of its base has a finite subcover. Suppose $\{u(x_i):i\in I\}$ is such a cover which has no finite subcover. Then for each finite subset I_0 of I,

$$\bigcup_{i\in I_0} u(x_i)\neq S(B)$$

and hence

$$\bigcap_{i\in I_0} u(x_i^*)=u\Big(\bigwedge_{i\in I_0} x_i^*\Big)\neq\emptyset=u(0).$$

Therefore

$$\bigwedge_{i\in I_0} x_i^*\neq 0.$$

Thus $\{x_i^*:i\in I\}$ has the fip and so can be extended to an ultrafilter U in B. Since, for each $i\in I$, $x_i^*\in U$,

$$U\notin\bigcup_{i\in I} u(x_i).$$

But this contradicts the fact that $\{u(x_i):i\in I\}$ covers $S(B)$. Therefore we can conclude that $S(B)$ is compact.

Since $u(x)=S(B)-u(x^*)$, each element of the base is clopen. Therefore $S(B)$ is a Boolean space.

By lemma 6.4, and the proof of theorem 6.1, u is an isomorphism of B onto $\boldsymbol{C}(S(B))=u[B]$.

There is an interesting result dual to Stone's; it asserts in a topological framework what Stone's result asserts algebraically.

THEOREM 6.6. *Each Boolean space is homeomorphic to the Stone space of its characteristic algebra.*

PROOF. Let X be a Boolean space. We show that the map $v:X\to S(\boldsymbol{C}(X))$ defined by

$$v(x)=\{A\in\boldsymbol{C}(X):x\in A\}$$

is a homeomorphism.

For each $A\in\boldsymbol{C}(X)$, either $x\in A$ or $x\in \boldsymbol{c}A$, and so $v(x)$ is an ultrafilter in $\boldsymbol{C}(X)$, that is, $v(x)\in S(\boldsymbol{C}(X))$.

v is a one–one map; since if x,y are distinct elements of X, because X is a Hausdorff space, there are disjoint elements A,A' of $\boldsymbol{C}(X)$ containing x and y respectively. Therefore $v(x)\neq v(y)$.

v maps X onto $S(\boldsymbol{C}(X))$; for suppose U is an ultrafilter in $\boldsymbol{C}(X)$. Then, in particular, U has the fip and so can be extended to an ultrafilter in $\langle \mathscr{P}(X), \subseteq \rangle$. Since X is compact this ultrafilter converges to some point $x\in X$ which is an adherent point of the ultrafilter. Therefore $x\in \bigcap\{\bar{A}: A\in U\}=\bigcap\{A: A\in U\}$ and so $U\subseteq v(x)$. Hence, because U is an ultrafilter, $U=v(x)$. This shows that v is onto.

Finally, because v is onto, for each A in $\boldsymbol{C}(X)$,

$$\{U\in S(\boldsymbol{C}(X)): A\in U\}=\{v(x): x\in A\},$$

so that v maps the base $\boldsymbol{C}(X)$ for the topology of X onto the base $\boldsymbol{C}(S(\boldsymbol{C}(X)))$ for the topology of $S(\boldsymbol{C}(X))$. It follows that v is a homeomorphism.

This last theorem shows that, in a sense, S and $\boldsymbol{C}$ are 'inverse' operations, and thus that there is a one–one correspondence between Boolean algebras on the one hand and Boolean spaces on the other.

§ 7. Atomless Boolean algebras and Cantor spaces

In ch. 9 we are going to prove a metamathematical result about atomless Boolean algebras, so in this section we explain what these are and we prove the algebraic result concerning them that we shall need later.

Let B be a Boolean algebra. An element $x\in B$ is said to be an *atom* if $x\neq 0$ and whenever $y\leqslant x$, then either $y=0$ or $y=x$. Thus x is an atom if there is no element of B between 0 and x, that is, if x is a minimal non-zero element of B.

EXERCISES

7.1. Show that $x\in B$ is an atom iff the principal filter generated by x is an ultrafilter.

7.2. Show that $x\in B$ is an atom iff there is exactly one ultrafilter in B which contains x.

A Boolean algebra B is said to be *atomic* if given any non-zero element $y\in B$, there is some atom $x\in B$ with $x\leqslant y$. For example, all power set Boolean algebras are atomic. In this case the atoms are just the one element sets.

EXERCISE 7.3. Suppose that B is a finite Boolean algebra. Prove that

B is atomic and is isomorphic to the power set Boolean algebra $\langle \mathscr{P}(A), \subseteq \rangle$, where A is the set of atoms of B. (Hint: define $\phi: B \to \mathscr{P}(A)$ by $\phi(y) = \{x \in A : x \leqslant y\}$ and show that ϕ is an isomorphism.)

A Boolean algebra is said to be *atomless* if it contains no atoms. One example of an atomless Boolean algebra is given in example 1.22. Another is given below.

We now ask the question: suppose that we know that the Boolean algebra B is atomless, what does this tell us about the Stone space $S(B)$ of B? The answer is provided by exercise 7.2 which informs us that x is an atom iff $u(x)$, the set of all ultrafilters containing x, has just one element. Thus B is atomless iff none of the basic open sets of $S(B)$ contains just one element and clearly this holds iff no open set in $S(B)$ has just one point in it. A topological space is said to be *perfect* if none of its one element subsets is open, or equivalently, if all its points are limit points. We have therefore just seen that

LEMMA 7.4. *A Boolean algebra B is atomless iff the Stone space of B is perfect.*

An easily proved result about perfect spaces that we shall need is

LEMMA 7.5. *If X is a perfect Hausdorff space, then no non-empty finite subset of X is open.*

PROOF. Suppose that X has a non-empty finite open subset. Let $G = \{x_0, \ldots, x_n\}$ be a non-empty finite open subset of X of least cardinal. Since X is perfect $n \geqslant 1$. Because X is a Hausdorff space for each $i < n$, there is an open subset U_i of X which contains x_i but not x_n. Then

$$\{x_0, \ldots, x_{n-1}\} = G \cap (U_0 \cup \cdots \cup U_{n-1})$$

is an open non-empty set containing fewer elements than G. This contradiction proves the lemma.

(Notice that this result is also true for perfect T_1-spaces.)

A *Cantor space* is a perfect compact Hausdorff space with a countable base of clopen sets. Thus a Cantor space is a perfect Boolean space with a countable base. The chief result of this section is

THEOREM 7.6. *If X is a Cantor space then X is homeomorphic to the space 2^ω with the product topology.*

PROOF. X has a countable base of clopen sets. Let $\langle \langle U_n, V_n \rangle : n \in \omega \rangle$ be an enumeration of all pairs of elements of this base that are disjoint.

By the previous lemma all non-empty open subsets of X are infinite and therefore it is easily seen that we can define, by recursion, for each $k\in\omega$ and each $f\in 2^k$, a non-empty clopen subset

$$X_{f(0)\dots f(k-1)}$$

of X such that

(a). $X=X_0\cup X_1$ and, in general,
$X_{f(0)\dots f(k-1)0}\cup X_{f(0)\dots f(k-1)1}=X_{f(0)\dots f(k-1)}$,

(b). $X_{f(0)\dots f(k-1)0}\cap X_{f(0)\dots f(k-1)1}=\emptyset$,

and

(c). if $U=U_k\cap X_{f(0)\dots f(k-1)}$ and $V=V_k\cap X_{f(0)\dots f(k-1)}$ are both non-empty then

$$U\subseteq X_{f(0)\dots f(k-1)0} \text{ and } V\subseteq X_{f(0)\dots f(k-1)1}.$$

Now, for each $f\in 2^\omega$, let

$$X_f=\bigcap_{k\in\omega} X_{f(0)\dots f(k-1)}.$$

X_f is the intersection of a collection of closed subsets of a compact space with the finite intersection property. It follows that X_f is not empty. We show that there is just one element in X_f.

Suppose not, and that a,b are distinct elements of X_f. X is a Hausdorff space and so we can find disjoint elements U_a, V_b of the base with $a\in U_a$ and $b\in V_b$. For some $n\in\omega$,

$$\langle U_a,V_b\rangle=\langle U_n,V_n\rangle.$$

Also

$$a\in U=U_a\cap X_{f(0)\dots f(n-1)}$$

and

$$b\in V=V_b\cap X_{f(0)\dots f(n-1)}.$$

Therefore, by (c),

$$U\subseteq X_{f(0)\dots f(n-1)0} \text{ and } V\subseteq X_{f(0)\dots f(n-1)1}.$$

Hence, by (b), if $f(n)=0$, then

$$b\notin X_{f(0)\dots f(n)}$$

and if $f(n)=1$, then

$$a\notin X_{f(0)\dots f(n)}.$$

But since

$$X_f\subseteq X_{f(0)\dots f(n)},$$

in either case this is a contradiction.

Thus for each $f\in 2^\omega$ there is a unique element, say x_f, in X_f. It is clear from (a) that for each $x\in X$ there is some $f\in 2^\omega$ such that $x=x_f$. So if we

define the map $\phi: X \to 2^\omega$ by

$$\phi(x_f) = f,$$

then ϕ is a one–one map of X onto 2^ω. To show that ϕ is continuous it will be sufficient to show that the inverse image of an arbitrary element of some sub-base for the topology of 2^ω is open in X. 2^ω has a sub-base consisting of all the sets

$$B_{ki} = \{f \in 2^\omega : f(k) = i\},$$

with $k \in \omega$ and $i \in \{0,1\}$.

For each $k \in \omega$, and $i \in \{0,1\}$, let

$$Y_{ki} = \bigcup_{f \in 2^k} X_{f(0)\ldots f(k-1)i}.$$

Then Y_{ki} is open and

$$x_f \in Y_{ki} \quad \text{iff} \quad f(k) = i \quad \text{iff} \quad f \in B_{ki}.$$

Hence $Y_{ki} = \phi^{-1}(B_{ki})$, which is therefore open.

This shows that ϕ is continuous. A one–one continuous mapping of a compact Hausdorff space onto another compact Hausdorff space is a homeomorphism, and so our proof is completed.

As a consequence we obtain the following algebraic result:

COROLLARY 7.7. *Any two countable atomless Boolean algebras are isomorphic.*

PROOF. Suppose B_0 and B_1 are countable atomless Boolean algebras. Then their Stone spaces are perfect Boolean spaces with countable bases, and so by the previous theorem they are homeomorphic. The characteristic algebras of these Stone spaces are therefore isomorphic, and so it follows from theorem 6.6 that B_0 and B_1 are isomorphic.

It is well known that the topological space 2^ω with the product topology is homeomorphic to the *Cantor discontinuum* (also called *Cantor's ternary set*). (See, e.g. KELLEY [1955] page 165.) This explains our terminology. It also shows that all countable atomless Boolean algebras are isomorphic to the Boolean algebra of clopen subsets of the Cantor discontinuum. This is our second example of an atomless Boolean algebra.

A Boolean algebra B is said to be *dense* if given any distinct elements x, y of B with $x \leqslant y$, there is an element z of B distinct from x and y with $x \leqslant z$ and $z \leqslant y$.

EXERCISE 7.8. Show that a Boolean algebra is dense iff it is atomless. Thus corollary 7.7 can be restated as: *any two countable dense Boolean*

algebras are isomorphic. This result should be compared with that of exercise 9.1.4.

Finally we remark that the results of this section show that if B is a countable atomless Boolean algebra then B contains precisely $2^{\aleph_0}$ distinct ultrafilters. This situation should be contrasted with that obtaining in theorem 6.1.5.

§8. Historical and bibliographical remarks

Boolean algebra is named after the English mathematician George Boole (1815–1864). Bertrand Russell once remarked that Boole discovered pure mathematics. This is an exaggeration, but it is true that Boole was one of the first to view mathematics as the study of abstract structure rather than as the science of magnitude. His *The Mathematical Analysis of Logic, being an essay towards a calculus of deductive reasoning* (BOOLE [1847]), was really the first successful attempt to apply mathematical techniques to logic. This short book is still in print and is well worth reading.

In recent years Boolean algebras have been studied a great deal by mathematicians and they have come to be of importance in digital computing. Of the many elementary textbooks dealing with Boolean algebras we mention only GOODSTEIN [1963]. More advanced works are BIRKHOFF [1967], HALMOS [1963] and SIKORSKI [1964].

The system of axioms L_1–L_5 is due to HUNTINGTON [1904] where the equivalence between complemented distributive lattices and Boolean algebras was first worked out. The ultrafilter theorem is due to TARSKI [1930]. Usually it is stated in terms of ideals and is then known as the prime ideal theorem. Tarski's lemma was first proved by RASIOWA and SIKORSKI [1951] using topological methods. The simple proof we have given is due to Tarski (see FEFERMAN [1952]).

The Stone representation theorem was first proved by STONE [1936] and theorem 6.2 by STONE [1937]. The introduction of filters and ultrafilters into topology is due to CARTAN [1937] and [1937i].

CHAPTER 2

PROPOSITIONAL CALCULUS

In this chapter we formulate the propositional calculus and we prove completeness and compactness theorems for it. Although no previous knowledge of propositional logic is necessary for an understanding of this chapter, our treatment is fairly brisk and touches only on the essentials. The reader desiring a more detailed survey should consult other textbooks.

§1. The system SC

It is clear that the argument:

Either *today is Wednesday* or MY NAME IS NAPOLEON BONAPARTE.

It is not the case that *today is Wednesday*.

Therefore MY NAME IS NAPOLEON BONAPARTE.

is logically valid and that its logical validity does not at all depend on the meaning of the words in italics or capitals. If we replaced them by any other two sentences we would still have a valid argument. The validity of the argument depends only on the logical words such as *or* and *not* that are used and not on the internal structure of the sentences they connect.

Propositional calculus is designed to formalize that part of logic where the validity of the argument depends only on how sentences are strung together and not on their internal structure. The system SC we use to formalize this part of logic is described in the following pages.

We begin by saying what the *basic symbols* of SC are. These fall into three categories:

(a). *Propositional variables.* The propositional variables of SC are the elements of the infinite set

$$P=\{p_n: n\in\omega\}.$$

(b). *Logical connectives.* These are

$\neg$ (read 'not') and $\wedge$ (read 'and').

(c). *Punctuation symbols.* These are

((left hand bracket) and) (right hand bracket).

A *string* of symbols is just a finite sequence of basic symbols. For example

$$)p_1 \neg (p_2$$
$$\neg (p_1 \wedge p_2)$$
$$((\neg p_1 p_2 \wedge$$

are all strings. We single out a subset of the strings. The elements of this subset are called *formulas*. Intuitively these are those strings which can be interpreted meaningfully and unambiguously when we interpret the propositional variables in them as propositions.

(Philosophically minded readers will note that we have blurred the distinction between sentences and propositions. This distinction is important, but not for the purposes of this book. The calculus we are describing is sometimes thought of as a calculus of sentences instead of as a calculus of propositions. In this case it is often called the sentential calculus. This explains our use of 'SC' for the system. We use it because we want to be able to use 'PC' for the predicate calculus.)

The set of formulas is defined recursively as follows:

(1). A string consisting of a single propositional variable by itself is a formula.

(2). If ϕ and ψ are formulas, so too are $\neg\phi$ and $(\phi \wedge \psi)$.

(3). A string is a formula only if it follows from a finite number of applications of rules (1) and (2) that it is a formula.

We let F be the set of formulas of SC. Thus F is the smallest set of strings which contains all the propositional variables by themselves, and whenever it contains ϕ and ψ also contains $\neg\phi$ and $(\phi \wedge \psi)$. For example, the following strings are all formulas:

(a). $\neg(p_0 \wedge \neg p_1)$

(b). $\neg(\neg p_0 \wedge \neg p_1)$

(c). $\neg(\neg(p_0 \wedge \neg p_1) \wedge \neg\neg(\neg(p_1 \wedge \neg p_2) \wedge \neg\neg(p_0 \wedge \neg p_2)))$.

For reasons of economy we have chosen to take only $\neg$ and $\wedge$ as basic logical connectives. In practice however we also use other connectives such as disjunction and implication. We introduce these as abbreviatory devices in the way explained below. An alternative procedure would have been to take these new symbols as basic symbols in addition to $\neg$ and $\wedge$. If we had done this we could have shown that they were superfluous by showing

that any formula involving them is equivalent, in a sense to be made precise later, to a formula containing only the connectives $\neg$ and $\wedge$.

The new connectives we introduce as abbreviatory devices are

$\vee$ (read 'or'),

$\rightarrow$ (read 'implies' or 'only if')

and

$\leftrightarrow$ (read 'iff').

They are used in the following way:

$(\phi \vee \psi)$ abbreviates $\neg(\neg\phi \wedge \neg\psi)$,

$(\phi \rightarrow \psi)$ abbreviates $\neg(\phi \wedge \neg\psi)$,

and

$(\phi \leftrightarrow \psi)$ abbreviates $((\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi))$,

that is, in full,

$$(\neg(\phi \wedge \neg\psi) \wedge \neg(\psi \wedge \neg\phi)).$$

In order to make formulas easier to read, we often omit brackets or use brackets of a different shape. Thus the formulas (a), (b) and (c) above will usually appear as

(a′). $p_0 \rightarrow p_1$

(b′). $p_0 \vee p_1$

(c′). $[p_0 \rightarrow p_1] \rightarrow [(p_1 \rightarrow p_2) \rightarrow (p_0 \rightarrow p_2)]$

which are easier to read than the originals.

The intuitive interpretation of formulas of SC will be clear from our earlier remarks, but now we want to make this more precise. We have an interpretation of a particular formula when we have interpreted the propositional variables occurring in it as particular propositions. But since what we are interested in is whether a formula is true or false under a given interpretation, all we really need to know is whether the proposition assigned to a particular propositional variable is *true* or *false*. Thus we can regard an interpretation as an assignment of the truth values *true* and *false* to the propositional variables. Formally it is convenient to represent these truth values by the elements 1 and 0 of the two element Boolean algebra **2**. We are now ready for the formal definition.

A *realization* of SC is a map of P, the set of propositional variables of SC, into the two element Boolean algebra **2**. Given a realization f of SC we extend it to F, the set of all formulas of SC, by the following recursive definition:

For any formulas ϕ,ψ, if $f(\phi)$ and $f(\psi)$ have already been defined, then

$$f(\phi \wedge \psi)=f(\phi)\wedge f(\psi)$$

and

$$f(\neg\phi)=f(\phi)^*.$$

It follows from this definition that, for example,

$$f(\phi \vee \psi)=f(\neg(\neg\phi \wedge \neg\psi))=(f(\phi)^* \wedge f(\psi)^*)^*=f(\phi)\vee f(\psi).$$

Note that on the right hand side of these expressions the symbols '$\wedge$', '$\vee$' are the meet and join operators of the Boolean algebra **2**, whereas on the left hand side they are symbols of SC. We continue to use them in this ambiguous way since it will always be clear from the context which is intended.

If the realization f, extended in the way just described, maps the formula ϕ onto 1, the maximum of **2**, we say that f *satisfies* ϕ and we write $f \vDash \phi$. Thus

$$f \vDash \phi \quad \text{iff} \quad f(\phi)=1.$$

A formula ϕ is said to be *satisfiable* if it is satisfied by some realization. A formula which is satisfied by all realizations is called a *tautology* or a *valid* formula. A set Σ of formulas is said to be *satisfiable* if there is a realization which simultaneously satisfies all the formulas in Σ.

EXERCISES

1.1. Show that the relation $f \vDash \phi$ can be defined recursively as follows:
(a). for $p_n \in P$, $f \vDash p_n$ iff $f(p_n)=1$,
(b). $f \vDash \neg\phi$ iff not $f \vDash \phi$,
(c). $f \vDash \phi \wedge \psi$ iff $f \vDash \phi$ and $f \vDash \psi$.

1.2. Show that ϕ is a tautology iff $\neg\phi$ is not satisfiable.

1.3. Show that

$$f(\phi\rightarrow\psi)=f(\neg\phi \vee \psi)=f(\phi)^* \vee f(\psi)$$

and hence that

$$f(\phi\rightarrow\psi)=0 \quad \text{iff} \quad f(\phi)=1 \quad \text{and} \quad f(\psi)=0.$$

1.4. Let $\Sigma=\{\phi_1,\ldots,\phi_n\}$ be a finite set of formulas. Prove that if every realization which satisfies $\Sigma \cup \{\phi\}$ also satisfies ψ, then $(\phi_1 \wedge \cdots \wedge \phi_n)\rightarrow(\phi\rightarrow\psi)$ is a tautology.

As we have already remarked, intuitively, $f(\phi)=1$ says that 'ϕ is true' and $f(\phi)=0$ says that 'ϕ is false'. Exercise 1.1 shows that, for example, $\phi \wedge \psi$

is 'true' iff both ϕ and ψ are 'true'; this agrees with our intuitive idea that '$\wedge$' represents 'and'. It can therefore be seen that our definition of 'tautology' is essentially the same as the usual definition by truth tables.

Suppose ϕ is a formula and that $p_0,\ldots,p_{n-1}$ are the only propositional variables that occur in ϕ. Clearly if f,g are realizations of SC such that for $i<n$, $f(p_i)=g(p_i)$, then also $f(\phi)=g(\phi)$. Thus ϕ determines a function, say v_ϕ, from $\mathbf{2}^n$ into $\mathbf{2}$. v_ϕ is the function given by

$$v_\phi(t_0,\ldots,t_{n-1})=f(\phi),$$

where f is any realization of SC which for $i<n$ maps p_i onto the element t_i of $\mathbf{2}$. The logical connectives $\neg$ and $\wedge$ together form a *complete set of connectives* in the sense that for each $n\in\omega$, $n\neq 0$, given any function $v:\mathbf{2}^n\to\mathbf{2}$ there is a formula ϕ built up from the propositional variables $p_0,\ldots,p_{n-1}$ using just the connectives $\neg$ and $\wedge$ such that $v=v_\phi$.

EXERCISE 1.5. Verify this last assertion.

It also follows from this that the formula ϕ is a tautology iff v_ϕ takes just the value 1. Since v_ϕ is a function of just a finite number of variables we can check in an effective manner whether it takes the value 0 or not. Thus we have a decision procedure for determining whether or not an arbitrary formula of SC is a tautology. In the case of predicate calculus, considered in the next chapter, no such decision procedure is available and we have to use an axiom system in order to generate those formulas true under all interpretations. By the remarks just made such an axiomatization is not necessary in the case of propositional calculus, but none the less it is interesting to consider axiomatizations for propositional calculus especially as it enables us to introduce most of the basic ideas involved in a particularly simple setting. This we now do.

We select a certain subset of the collection of tautologies whose elements will be our basic valid formulas or *axioms*, and we use *rules of inference* to enable us to derive new valid formulas from the axioms. Our axiomatization will be adequate if all valid formulas are derivable in this way.

As *axioms* we take all formulas of one of the forms

SC1. $\phi\to(\psi\to\phi)$.

SC2. $[\phi\to(\psi\to\chi)]\to[(\phi\to\psi)\to(\phi\to\chi)]$.

SC3. $(\neg\phi\to\neg\psi)\to(\psi\to\phi)$.

SC4a. $(\phi\wedge\psi)\to\phi$.

SC4b. $(\phi\wedge\psi)\to\psi$.

SC5. $[\chi\to\phi]\to[(\chi\to\psi)\to(\chi\to[\phi\wedge\psi])]$.
SC6a. $\phi\to(\phi\vee\psi)$.
SC6b. $\psi\to(\phi\vee\psi)$.
SC7. $[\phi\to\chi]\to[(\psi\to\chi)\to([\phi\vee\psi]\to\chi)]$.

where ϕ, ψ and χ are any formulas of SC. The reason for this particular choice of axioms will become apparent in the sequel.

The only *rule of inference* that we have is *modus ponens:* ψ is an *immediate consequence* of ϕ and $\phi\to\psi$, where ϕ and ψ are any formulas of SC. A *proof* of a formula ϕ is a finite sequence $\phi_1,\ldots,\phi_n$ of formulas such that $\phi_n=\phi$ and for each $k\leqslant n$, ϕ_k is either an axiom or for some $i,j<k$ is an immediate consequence of ϕ_i and ϕ_j according to the rule modus ponens. A formula ϕ is said to be *provable* if there exists a proof of it. If ϕ is provable we write $\vdash\phi$.

LEMMA 1.6. *For each formula* ϕ *of* SC, $\vdash\phi\to\phi$.

PROOF. Consider the sequence of formulas

$$[\phi\to([\phi\to\phi]\to\phi)]\to[(\phi\to[\phi\to\phi])\to(\phi\to\phi)]$$
$$\phi\to([\phi\to\phi]\to\phi)$$
$$(\phi\to[\phi\to\phi])\to(\phi\to\phi)$$
$$\phi\to[\phi\to\phi]$$
$$\phi\to\phi$$

The first formula is an axiom of the form SC2, the second an axiom of the form SC1, the third is an immediate consequence of the first two, the fourth is an axiom of the form SC1 and the fifth is an immediate consequence of the third and fourth formulas. Thus the sequence is a proof of $\phi\to\phi$.

In future, as is usual, instead of exhibiting a sequence of formulas and then adding an explanation of why the sequence is a proof we will give each formula a number and put annotations at the side explaining how each line arises.

LEMMA 1.7. *For any formulas* ϕ,ψ *and* χ *of* SC

$$\vdash(\psi\to\chi)\to[(\phi\to\psi)\to(\phi\to\chi)].$$

PROOF. Consider the following sequence:

1. $[\phi\to(\psi\to\chi)]\to[(\phi\to\psi)\to(\phi\to\chi)]$, axiom SC2.
2. $([\phi\to(\psi\to\chi)]\to[(\phi\to\psi)\to(\phi\to\chi)])$
$\to((\psi\to\chi)\to([\phi\to(\psi\to\chi)]\to[(\phi\to\psi)\to(\phi\to\chi)]))$, axiom SC1.

3. $(\psi\to\chi)\to([\phi\to(\psi\to\chi)]\to[(\phi\to\psi)\to(\phi\to\chi)])$, modus ponens 1,2.
4. $((\psi\to\chi)\to([\phi\to(\psi\to\chi)]\to[(\phi\to\psi)\to(\phi\to\chi)]))$
$\to([(\psi\to\chi)\to[\phi\to(\psi\to\chi)]]\to[(\psi\to\chi)\to[(\phi\to\psi)\to(\phi\to\chi)]])$, axiom SC2.
5. $[(\psi\to\chi)\to[\phi\to(\psi\to\chi)]]\to[(\psi\to\chi)\to[(\phi\to\psi)\to(\phi\to\chi)]]$, modus ponens 3,4.
6. $(\psi\to\chi)\to[\phi\to(\psi\to\chi)]$, axiom SC1.
7. $(\psi\to\chi)\to[(\phi\to\psi)\to(\phi\to\chi)]$, modus ponens 5,6.

By the annotations this sequence is a proof of the last formula.

In order to shorten proofs, when we have shown that a given formula is provable and we want to use this formula in a later proof instead of repeating the proof we just give a reference to a place where it can be found. We do this in the next lemma.

LEMMA 1.8. *For any formulas* ϕ,ψ *of* SC,

$$\vdash \neg\phi\to(\phi\to\psi).$$

PROOF. We have the following proof sequence:

1. $(\neg\psi\to\neg\phi)\to(\phi\to\psi)$, axiom SC3.
2. $[(\neg\psi\to\neg\phi)\to(\phi\to\psi)]\to([\neg\phi\to(\neg\psi\to\neg\phi)]\to[\neg\phi\to(\phi\to\psi)])$, lemma 1.7.
3. $[\neg\phi\to(\neg\psi\to\neg\phi)]\to[\neg\phi\to(\phi\to\psi)]$, modus ponens 1,2.
4. $\neg\phi\to(\neg\psi\to\neg\phi)$, axiom SC1.
5. $\neg\phi\to(\phi\to\psi)$, modus ponens 3,4.

It must be remembered that we are using the connectives $\to$ and $\vee$ as abbreviatory devices. Thus lemma 1.6 tells us that

$$\vdash \neg(\phi\wedge\neg\phi)$$

for any formula ϕ. In particular, writing $\neg\phi$ for ϕ,

$$\vdash \neg(\neg\phi\wedge\neg\neg\phi),$$

hence

$$\vdash \phi\vee\neg\phi.$$

EXERCISE 1.9. Show that for any formulas ϕ,ψ,θ and χ:

(a). $\vdash(\phi\wedge\psi)\to(\psi\wedge\phi)$,
(b). $\vdash\phi\to[\psi\to(\phi\wedge\psi)]$,
(c). $\vdash[(\phi\wedge\chi)\vee(\psi\wedge\chi)]\to[(\phi\vee\psi)\wedge\chi]$,
(d). $\vdash(\chi\to\theta)\to[(\phi\to(\psi\to\chi))\to(\phi\to(\psi\to\theta))]$,

(e). $\vdash(\phi \vee \psi) \rightarrow(\chi \rightarrow[(\phi \wedge \chi) \vee(\psi \wedge \chi)])$,
(f). $\vdash[\phi \rightarrow(\psi \rightarrow \chi)] \rightarrow[(\phi \wedge \psi) \rightarrow \chi]$,
(g). $\vdash[(\phi \vee \psi) \wedge \chi] \rightarrow[(\phi \wedge \chi) \vee(\psi \wedge \chi)]$.

Let Σ be a set of formulas of SC. A *proof of the formula ϕ from the assumptions Σ* is a finite sequence of formulas $\phi_1,\ldots,\phi_n$ such that $\phi=\phi_n$ and for all $k \leqslant n$, ϕ_k is either an axiom, one of the formulas in Σ, or for some $i, j<k$ is an immediate consequence of ϕ_i and ϕ_j. If there is a proof of ϕ from the assumptions Σ we write $\Sigma \vdash \phi$. We allow the case in which Σ is infinite, but the next result shows that this adds nothing new.

THEOREM 1.10. THE FINITENESS THEOREM FOR PROPOSITIONAL CALCULUS *If $\Sigma \vdash \phi$, there is some finite subset Σ_0 of Σ such that $\Sigma_0 \vdash \phi$.*

PROOF. Let $\phi_1,\ldots,\phi_n$ be a proof of ϕ from the assumptions Σ. Let Σ_0 be the set of those formulas of Σ occurring in this proof. Σ_0 is finite and $\phi_1,\ldots,\phi_n$ is a proof of ϕ from the assumptions Σ_0.

Let *Taut* be the set of tautologies of SC and let *Prov* be the set of provable formulas. We now investigate the relationship between these two sets of formulas.

THEOREM 1.11. *Every provable formula of* SC *is a tautology.*

PROOF. The proof is by induction on the length of proofs. We show that each axiom is a tautology and that if ϕ and $\phi \rightarrow \psi$ are tautologies then so also is ψ. It will follow that a proof consists of a sequence of tautologies and, in particular, that every provable formula is a tautology.

We first show that any instance of the axiom schema

SC1. $\phi \rightarrow(\psi \rightarrow \phi)$

is a tautology. Let f be any realization of SC. Then

$$\begin{aligned} f(\phi \rightarrow(\psi \rightarrow \phi)) &= f(\neg \phi \vee(\neg \psi \vee \phi)) \\ &= f(\neg \phi) \vee f(\neg \psi) \vee f(\phi) \\ &= f(\phi)^* \vee f(\psi)^* \vee f(\phi) \\ &= 1 \vee f(\psi)^* \\ &= 1 . \end{aligned}$$

Thus for any formulas ϕ, ψ the formula $\phi \rightarrow(\psi \rightarrow \phi)$ is a tautology. The other axiom schemas can be checked in a similar way. We leave this as an exercise for the reader.

Now suppose that ϕ and $\phi \to \psi$ are tautologies and that f is any realization of SC. It follows that $f(\phi) = f(\phi \to \psi) = 1$, and so

$$\begin{aligned} f(\psi) &= [f(\phi) \wedge f(\phi)^*] \vee f(\psi) \\ &= [f(\phi) \vee f(\psi)] \wedge [f(\phi)^* \vee f(\psi)] \\ &= [1 \vee f(\psi)] \wedge f(\phi \to \psi) \\ &= 1 \wedge 1 \\ &= 1. \end{aligned}$$

Therefore ψ is also a tautology. This completes the proof.

COROLLARY 1.12. *There is no formula ϕ such that both ϕ and $\neg\phi$ are provable.*

PROOF. A realization f satisfies ϕ iff it does not satisfy $\neg\phi$. Therefore ϕ and $\neg\phi$ cannot both be tautologies. The result is now an immediate consequence of the previous theorem.

EXERCISES

1.13. Show that $(\phi \to \psi) \to ((\phi \to \neg\psi) \to \neg\phi)$ and all instances of the schemas SC2–SC7 are tautologies.

1.14. Prove that if $\Sigma \vdash \phi$ then any realization which satisfies each formula of Σ also satisfies ϕ.

Theorem 1.11 shows that *Prov* $\subseteq$ *Taut*. The next two sections are devoted to determining whether the converse inclusion also holds.

§2. The Lindenbaum algebra

We are now ready to apply our knowledge of Boolean algebras to the study of propositional calculus. To do this we must first impose the structure of a Boolean algebra on the set F of formulas of SC. The natural first idea is to define the relation $\preccurlyeq$ on F by

$$\phi \preccurlyeq \psi \quad \text{iff} \quad \vdash \phi \to \psi.$$

But although $\preccurlyeq$ is reflexive and transitive, it fails to be a partial ordering because it is not an antisymmetric relation. To see this, notice that by exercise 1.9(a), if ϕ, ψ are any two distinct formulas from F

$$\phi \wedge \psi \preccurlyeq \psi \wedge \phi \quad \text{and} \quad \psi \wedge \phi \preccurlyeq \phi \wedge \psi,$$

but $\phi \wedge \psi$ and $\psi \wedge \phi$ are distinct formulas. However, in an obvious sense,

they are equivalent, so we can overcome this difficulty by first defining the relation $\equiv$ on F by

$$\phi \equiv \psi \quad \text{iff} \quad \vdash \phi \to \psi \quad \text{and} \quad \vdash \psi \to \phi .$$

By lemma 1.6, $\equiv$ is a reflexive relation. It is clearly a symmetric relation and it is an immediate consequence of lemma 1.7 that it is a transitive relation. Thus $\equiv$ is an equivalence relation on F.

EXERCISE 2.1. Show that if $\phi \equiv \phi'$ and $\psi \equiv \psi'$, then

$$\vdash \phi \to \psi \quad \text{iff} \quad \vdash \phi' \to \psi' .$$

If ϕ is a formula in F, we let $|\phi|$ be the equivalence class under the relation $\equiv$ to which it belongs. Thus

$$|\phi| = \{\psi \in F : \phi \equiv \psi\} .$$

We let

$$F/\equiv \; = \{|\phi| : \phi \in F\} .$$

Now we can define the relation $\leqslant$ on $F/\equiv$ by

$$|\phi| \leqslant |\psi| \quad \text{iff} \quad \vdash \phi \to \psi .$$

Exercise 2.1 shows that this is an unambiguous definition.

That we have now overcome our earlier difficulty is shown by:

THEOREM 2.2. *$\mathscr{A} = \langle F/\equiv, \leqslant \rangle$ is a complemented distributive lattice, that is, a Boolean algebra. In this Boolean algebra:*

$$|\phi| = 1 \text{ iff } \vdash \phi \quad \text{and} \quad |\phi| = 0 \text{ iff } \vdash \neg\phi .$$

PROOF. By lemma 1.6, $\leqslant$ is a reflexive relation; by definition it is an antisymmetric relation and lemma 1.7 shows that it is a transitive relation. Thus $\leqslant$ is a partial ordering of $F/\equiv$.

Let $\{|\phi|, |\psi|\}$ be any two element subset of $F/\equiv$. We show that $|\phi \wedge \psi|$ is its infimum. By SC4a and SC4b

$$\vdash (\phi \wedge \psi) \to \phi \quad \text{and} \quad \vdash (\phi \wedge \psi) \to \psi$$

whence $|\phi \wedge \psi| \leqslant |\phi|$ and $|\phi \wedge \psi| \leqslant |\psi|$. So $|\phi \wedge \psi|$ is a lower bound for $\{|\phi|,|\psi|\}$. Suppose $|\chi|$ is a lower bound for $\{|\phi|,|\psi|\}$. Then $\vdash \chi \to \phi$ and $\vdash \chi \to \psi$. Hence by SC5 and two applications of modus ponens, $\vdash \chi \to (\phi \wedge \psi)$. This shows that $|\phi \wedge \psi|$ is the greatest lower bound, i.e. the infimum of $\{|\phi|,|\psi|\}$. Similarly $|\phi \vee \psi|$ is the supremum of $\{|\phi|,|\psi|\}$. We have now shown that $\mathscr{A}$ is a lattice.

By exercise 1.9(c) and (g) for any formulas ϕ,ψ,χ

$$|(\phi \vee \psi) \wedge \chi| = |(\phi \wedge \chi) \vee (\psi \wedge \chi)|,$$

hence, by the remarks above,

$$(|\phi| \vee |\psi|) \wedge |\chi| = (|\phi| \wedge |\chi|) \vee (|\psi| \wedge |\chi|),$$

and, by exercise 1.1.15, this shows that $\mathscr{A}$ is a distributive lattice.

By SC1, if $\vdash \phi$, then for any formula ψ, $\vdash \psi \to \phi$ and so $|\psi| \leqslant |\phi|$. Hence $|\phi| = 1$, the maximum element of $\mathscr{A}$. Similarly, using lemma 1.8, if $\vdash \neg\phi$, then $|\phi| = 0$, the minimum element of $\mathscr{A}$. Conversely if $|\phi| = 1$, then for any formula ψ, $|\psi| \leqslant |\phi|$ and so $\vdash \psi \to \phi$. Therefore choosing ψ so that $\vdash \psi$, by modus ponens $\vdash \phi$. Similarly, if $|\phi| = 0$ then $\vdash \neg\phi$.

Finally notice that we have already remarked that for any formula ϕ,

$$\vdash \phi \vee \neg\phi \quad \text{and} \quad \vdash \neg(\phi \wedge \neg\phi),$$

hence

$$|\phi| \vee |\neg\phi| = 1 \quad \text{and} \quad |\phi| \wedge |\neg\phi| = 0.$$

This shows that $\mathscr{A}$ is a complemented lattice, and that the complement of $|\phi|$ is $|\neg\phi|$.

We have now shown that $\mathscr{A}$ is a complemented distributive lattice, that is, a Boolean algebra, and our proof is completed.

Since, by corollary 1.12, not every formula of F is provable, $\mathscr{A}$ contains at least two elements. We call $\mathscr{A}$ the *Lindenbaum algebra* of SC.

§ 3. The completeness of propositional calculus

We are now going to show that all tautologies are provable. Together with theorem 1.11 this will imply that

$$Taut = Prov$$

and hence that our axiom system does just what we want it to do. In other words it will show that our axiomatization is *complete.*

LEMMA 3.1. *Let h be a homomorphism of the Lindenbaum algebra $\mathscr{A}$ into the Boolean algebra* **2**. *If f is defined on P by $f(p_n) = h(|p_n|)$, then f is a realization of* SC *such that for each formula ϕ of F*

$$f(\phi) = h(|\phi|).$$

PROOF. The proof is by induction on the number of logical connectives

in ϕ. By hypothesis it is true for propositional variables standing by themselves. Suppose it is true for the formulas ϕ and ψ. Then since h is a homomorphism

$$f(\phi \wedge \psi) = f(\phi) \wedge f(\psi) = h(|\phi|) \wedge h(|\psi|) = h(|\phi| \wedge |\psi|) = h(|\phi \wedge \psi|),$$

and

$$f(\neg\phi) = f(\phi)^* = h(|\phi|)^* = h(|\phi|^*) = h(|\neg\phi|),$$

and hence the result is true for all formulas.

The realization f is called the realization *induced* by h.

Now we observe that there is a natural correspondence between ultrafilters in the Boolean algebra $\mathscr{A}$ and realizations of SC. If U is an ultrafilter in $\mathscr{A}$, then, by theorem 1.4.9, the quotient algebra $\mathscr{A}/U$ is isomorphic to the two element Boolean algebra **2**. Thus corresponding to U we obtain the realization f_U of SC induced by the canonical homomorphism of $\mathscr{A}$ into $\mathscr{A}/U$. On the other hand if f is a realization of SC let

$$U_f = \{|\phi| : f(\phi) = 1\}.$$

EXERCISE 3.2. Show that U_f as defined above is an ultrafilter in $\mathscr{A}$.

This exercise shows that corresponding to each realization of SC we obtain an ultrafilter in $\mathscr{A}$. This correspondence is the essential idea in the proof of the theorem that follows.

THEOREM 3.3. THE COMPLETENESS THEOREM FOR PROPOSITIONAL CALCULUS *Each tautology of* SC *is provable.*

PROOF. Suppose that ϕ is an unprovable formula of SC. Then in the Lindenbaum algebra $\mathscr{A}$, $|\phi| \neq 1$ and so $|\neg\phi| \neq 0$. Therefore, by corollary 1.3.6 to the ultrafilter theorem, there is an ultrafilter U in $\mathscr{A}$ which contains $|\neg\phi|$. By theorem 1.4.9, $\mathscr{A}/U \simeq \mathbf{2}$. Let f be the realization of SC induced by the canonical homomorphism h of $\mathscr{A}$ into $\mathscr{A}/U$. Since $|\neg\phi| \in U$, $h(|\neg\phi|) = 1$ and so $f(\neg\phi) = 1$. It follows that $f(\phi) = 0$ and so ϕ is not a tautology.

This concludes the proof.

THEOREM 3.4. THE DEDUCTION THEOREM FOR PROPOSITIONAL CALCULUS *Let Σ be a set of formulas of* SC *and let ϕ, ψ be any formulas of* SC. *If $\Sigma \cup \{\phi\} \vdash \psi$, then $\Sigma \vdash \phi \rightarrow \psi$.*

PROOF. Suppose $\Sigma \cup \{\phi\} \vdash \psi$. By the finiteness theorem there is some finite subset $\{\phi_1, \ldots, \phi_n\}$ of Σ such that

$$\{\phi_1, \ldots, \phi_n\} \cup \{\phi\} \vdash \psi.$$

Therefore, by exercise 1.14, every realization satisfying $\{\phi_1,\ldots,\phi_n\}\cup\{\phi\}$ also satisfies ψ. Hence, by exercise 1.4,

$$(\phi_1\wedge\cdots\wedge\phi_n)\rightarrow(\phi\rightarrow\psi)$$

is a tautology and so, by the completeness theorem, is provable. Clearly therefore

$$\{\phi_1,\ldots,\phi_n\}\vdash\phi\rightarrow\psi,$$

and hence

$$\Sigma\vdash\phi\rightarrow\psi.$$

The deduction theorem is usually proved syntactically by giving a procedure for converting a proof of ψ from the assumptions $\Sigma\cup\{\phi\}$ into a proof of $\phi\rightarrow\psi$ from the assumptions Σ. In some ways a syntactical proof is preferable since it gives more information than a semantic proof. Thus an analysis of the usual syntactic proof of the deduction theorem (see e.g. MENDELSON [1964] page 32) shows that given any system of propositional calculus, whose only rule of inference is modus ponens, then the deduction theorem holds for the system iff all formulas of the forms of the axiom schemas SC1 and SC2 are provable in the system.

In order to prove the completeness theorem we needed to show that certain formulas are provable. It will be evident by now that we chose our axioms so as to make this easy. In fact it is possible to show that the axiom schemas SC1–SC3 alone form a complete axiomatization for SC (see CHURCH [1956] pages 118–128). Therefore it is possible to prove all the axioms of the forms SC4–SC7 from the axioms of the forms SC1–SC3. We have refrained from doing this since finding proofs of formulas in SC is somewhat tedious and the proofs when they are found are apt to be long, even for quite simple formulas. This can be seen by examining the proofs of lemmas 1.7 or 1.8 or by trying to solve exercise 1.9. It should however be noted that we made things rather difficult for ourselves by postponing the proof of the deduction theorem until we could give a semantic proof of it. The use of this theorem, as well as of other derived rules, makes it much easier to prove that particular formulas are provable. We do not pursue this point since now that we know that the completeness theorem holds an even easier method is available to us for doing this.

§ 4. The compactness of propositional calculus

A set Σ of formulas of SC is said to be *consistent* if there is no formula ϕ

such that both ϕ and $\neg\phi$ are provable from the assumptions Σ. Corollary 1.12 shows that the empty set of formulas is consistent. We say that a single formula ϕ is *consistent* if the set $\{\phi\}$ is consistent.

LEMMA 4.1. *If Σ is a set of formulas and ϕ is any formula then $\Sigma \cup \{\phi\}$ is consistent iff not $\Sigma \vdash \neg\phi$.*

PROOF. If $\Sigma \vdash \neg\phi$ then obviously $\Sigma \cup \{\phi\} \vdash \neg\phi$; but clearly $\Sigma \cup \{\phi\} \vdash \phi$ and therefore $\Sigma \cup \{\phi\}$ is not consistent.

Suppose that $\Sigma \cup \{\phi\}$ is not consistent. Say $\Sigma \cup \{\phi\} \vdash \psi$ and $\Sigma \cup \{\phi\} \vdash \neg\psi$. Then, by the deduction theorem, $\Sigma \vdash \phi \to \psi$ and $\Sigma \vdash \phi \to \neg\psi$. By exercise 1.13

$$(\phi \to \psi) \to [(\phi \to \neg\psi) \to \neg\phi]$$

is a tautology and hence, by the completeness theorem, provable. Therefore, using modus ponens twice, $\Sigma \vdash \neg\phi$.

It follows from this lemma that ϕ is consistent iff $\neg\phi$ is not provable. The completeness theorem is equivalent to the statement that $\neg\phi$ is not provable iff ϕ is satisfiable. Thus the completeness theorem can be restated as

A formula ϕ is consistent iff it is satisfiable.

Clearly a finite set $\{\phi_1, \ldots, \phi_n\}$ of formulas is consistent iff the single formula $(\phi_1 \wedge \ldots \wedge \phi_n)$ is consistent. So the completeness theorem can also be stated as

A finite set of formulas is consistent iff it is satisfiable.

We now consider the question as to whether this remains true if we replace 'finite' by 'infinite'. That it does is a consequence of the next result.

THEOREM 4.2. THE COMPACTNESS THEOREM FOR PROPOSITIONAL CALCULUS *Let Σ be a set of formulas of SC. Then Σ is satisfiable iff every finite subset of Σ is satisfiable.*

PROOF. Clearly if Σ is satisfiable then so also is each finite subset of Σ.

For the converse suppose that each finite subset of Σ is satisfiable. Suppose $n < \omega$ and that we have found a map f of $\{p_i : i < n\}$ into **2** such that for each finite subset Σ_0 of Σ there is a realization which satisfies Σ_0 and agrees with f on the set $\{p_i : i < n\}$. (In the case $n = 0$ this is equivalent to our hypothesis that each finite subset of Σ is satisfiable.) Then we show that f can be extended to the set $\{p_i : i \leqslant n\}$ so that this same property holds.

Suppose that this property does not hold if we put $f(p_n) = 0$. Then it must be the case that there is some finite subset Σ_0 of Σ which is not satisfied by any realization which agrees with f on $\{p_i : i < n\}$ and in which p_n gets the value 0. Let Σ_1 be any finite subset of Σ. Then $\Sigma_2 = \Sigma_0 \cup \Sigma_1$ is a finite subset

of Σ and so, by hypothesis, is satisfied by a realization which agrees with f on $\{p_i : i<n\}$. By the choice of Σ_0 in this realization p_n must be mapped onto 1. Therefore Σ_1 is satisfied by a realization which agrees with f on $\{p_i : i<n\}$ and in which p_n gets the value 1.

It follows that we can define, by recursion on n, a realization f of SC such that for each $n<\omega$, every finite subset of Σ is satisfied by a realization which agrees with f on $\{p_i : i<n\}$.

f satisfies Σ. For suppose $\phi\in\Sigma$, then by choosing n large enough so that all the propositional variables of ϕ occur in $\{p_i : i<n\}$, it can be seen that f satisfies ϕ.

This completes the proof.

Let α be an infinite cardinal. If we had formulated SC so that it had $\{p_\xi : \xi<\alpha\}$ as its set of propositional variables the compactness theorem would still hold and we could prove it in essentially the same way. We also have the following topological proof:

We give the set $\{0,1\}$ the discrete topology. Since each formula ϕ contains only finitely many propositional variables the set $\bar{\phi}$ of those realizations which satisfy ϕ is open in $\{0,1\}^{P'}$ with the product topology, where $P'=\{p_\xi : \xi<\alpha\}$. Since those realizations which satisfy ϕ are precisely those which do not satisfy $\neg\phi$, $\bar{\phi}$ is also closed. Let $T=\{\bar{\phi} : \phi\in\Sigma\}$. By hypothesis T has the finite intersection property, and $\{0,1\}^{P'}$ being the product of compact spaces is also compact. Therefore $\bigcap\{\bar{\phi} : \phi\in\Sigma\}$ is not empty. But $\bigcap\{\bar{\phi} : \phi\in\Sigma\}$ is the set of those realizations which satisfy Σ.

The fact that the compactness theorem is equivalent to the compactness of the space $\{0,1\}^{P'}$ explains how it gets its name. The syntactic analogue of the compactness theorem is a straightforward consequence of the finiteness theorem.

THEOREM 4.3. *A set Σ of formulas is consistent iff each finite subset of Σ is consistent.*

PROOF. Suppose Σ has an inconsistent finite subset Σ_0. Any formula provable from the assumptions Σ_0 is also provable from the assumptions Σ and therefore Σ is also inconsistent.

Conversely suppose that Σ is not consistent. Then for some formula ϕ, $\Sigma\vdash\phi$ and $\Sigma\vdash\neg\phi$. By the finiteness theorem there are finite subsets Σ_1,Σ_2 of Σ such that $\Sigma_1\vdash\phi$ and $\Sigma_2\vdash\neg\phi$. Then $\Sigma_0=\Sigma_1\cup\Sigma_2$ is an inconsistent finite subset of Σ since both ϕ and $\neg\phi$ are provable from the assumptions Σ_0.

Together the last two theorems yield the following extension of the completeness theorem.

THEOREM 4.4. THE GENERALIZED COMPLETENESS THEOREM FOR PROPOSITIONAL CALCULUS

Σ is a consistent set of sentences of SC *iff Σ is satisfiable.*

PROOF. By theorem 4.3, Σ is consistent iff each finite subset of Σ is consistent, and by the completeness theorem this holds iff each finite subset of Σ is satisfiable. The compactness theorem says that this is true iff Σ is satisfiable and so our proof is finished.

As an application of the compactness theorem for propositional calculus we give the following amusing example.

THEOREM 4.5. *Let B be an infinite set of boys each of whom has at most a finite number of girl friends. If for each integer k, any k of the boys have between them at least k girl friends, then it is possible for each boy to marry one of his girl friends without any of them committing bigamy.*

PROOF. We shall need to use the following lemma dealing with the finite case.

LEMMA. *If C is a set of m boys and for each $k \leqslant m$, any k of the boys have at least k girl friends between them, then it is possible for each boy to marry one of his girl friends without any of them committing bigamy.*

PROOF OF LEMMA. The proof is by induction on m. The result is obvious in case $m=1$. We assume it holds whenever $m<n$ and that there are n boys in C. There are two cases to consider:

(a). If $k<n$, any k boys have at least $k+1$ girl friends between them.

(b). For some $k<n$, there is a set S of k boys who have exactly k girl friends between them.

In case (a) we marry one of the boys to any one of his girl friends. We can then apply the induction hypothesis to the remaining boys and girls.

In case (b) we use the induction hypothesis to marry each boy in S to a girl friend so that bigamy is not committed. We claim that for $l \leqslant n-k$, any l of the remaining boys have between them at least l girl friends among the remaining girls. For if T were a set of l remaining boys having fewer than l girl friends left, then $S \cup T$ would be a set of $k+l$ boys having fewer than $k+l$ girl friends altogether, contradicting the original hypothesis. It follows that we can again apply the induction hypothesis to the remaining boys and girls. This proves the lemma.

We are now ready to prove the main theorem. Let $B=\{b_i : i \in I\}$ and let $G=\{g_j : j \in J\}$ be the collection of all girl friends of the boys in B. We let L

be the propositional language whose propositional variables are the elements of the set

$$\{p_{ij}: i \in I \quad \text{and} \quad j \in J\}.$$

Let Σ be the set of formulas of L consisting of:

(1). For each $i \in I$, the formula

$$p_{ij_0'} \vee \cdots \vee p_{ij_{k'}},$$

where $g_{j_0'}, \ldots, g_{j_{k'}}$ are all the girl friends of b_i.

(2). For each $i \in I$ and each distinct pair j, j' of elements of J, the formula

$$\neg(p_{ij} \wedge p_{ij'}).$$

(3). For each $j \in J$ and each distinct pair i, i' of elements of I, the formula

$$\neg(p_{ij} \wedge p_{i'j}).$$

Let Σ_0 be a finite subset of Σ. We show that Σ_0 is satisfiable. Let C be the set of all those boys b_i such that for some $j \in J$ the propositional variable p_{ij} occurs in some formula of Σ_0. C is a finite set of boys to which the hypothesis of the lemma applies. Hence each boy in C can marry one of his girl friends without bigamy being committed. We assign p_{ij} the value 1 iff b_i marries g_j under this arrangement, and 0 otherwise. It is clear that the realization so defined satisfies Σ_0.

We can therefore conclude from the compactness theorem that there is some realization, say f, satisfying Σ. We now marry b_i to g_j iff $f(p_{ij}) = 1$. Since Σ contains all the formulas of (1) each boy marries one of his girl friends. Bigamy is not committed because Σ contains all the formulas of (2) and (3).

This completes the proof of the theorem.

It should be noted that the theorem is false if we drop the restriction that each boy has only a finite number of girl friends (which is, after all, a plausible assumption). We leave it to the reader to construct a counter-example in this case. (This is connected with the fact that the compactness theorem fails for propositional languages with infinitely long formulas. This is discussed further in ch. 14.)

EXERCISE 4.6. It is easy to see that a partial ordering on a finite set can be extended to a total ordering. Use the compactness theorem to show that this result is also true for infinite sets.

EXERCISE 4.7. A *graph* is an ordered pair $\langle G,R\rangle$ where R is a symmetric relation on G. Distinct elements g,g' of G are said to be *connected* in the graph if $\langle g,g'\rangle \in R$. A graph is *k-chromatic* if G can be partitioned into k disjoint sets so that connected points are not in the same set. Use the compactness theorem to show that a graph is k-chromatic iff all its finite subgraphs are k-chromatic.

§5. Historical and bibliographical remarks.

For a general account of the history of modern logic the reader is referred to the comprehensive work of WILLIAM and MARTHA KNEALE [1962] and for a survey of the early literature to the historical notes in CHURCH [1956] pages 155–166.

The algebra $\mathscr{A}$ is called the Lindenbaum algebra of SC in memory of the Polish mathematician Adolf Lindenbaum who was killed by the Nazis in the summer of 1941. Lindenbaum and Tarski developed the idea of the algebra independently around 1935; Lindenbaum's results were not published. The justice of calling $\mathscr{A}$ after Lindenbaum is discussed in RASIOWA and SIKORSKI [1963] (footnote to page 245).

The completeness of propositional calculus was first proved by POST [1921] using conjunctive normal forms. Another proof was given by KALMAR [1934–35]. Kalmar's method is the one used in most of the standard textbooks.

The proofs of Post and Kalmar both provide explicit recipes for constructing a proof of a given tautology. The proof that we have given, which is due to RASIOWA and SIKORSKI [1951], does not have this character, and since it depends on the ultrafilter theorem is not a constructive proof. However when we come to consider the predicate calculus in the next chapter we shall see that although there is an analogous completeness theorem this result cannot be proved constructively. In this case the Boolean algebra proof has a great advantage in terms of elegance and the way it highlights the algebraic nature of the result. The use of Boolean methods to prove the completeness of propositional calculus serves not only as an introduction to the ideas involved but also helps to show the similarities, and the differences, between the two cases.

The proof of the compactness theorem for propositional calculus is based on a combinatorial lemma of R. RADO [1949]. We have taken theorem 4.5 from HALMOS and VAUGHAN [1950] where the history of the result is described.

CHAPTER 3

PREDICATE CALCULUS

In this chapter we formulate the predicate calculus with equality and we prove that the axiom system that we give is complete. No previous knowledge of predicate logic is necessary for an understanding of this chapter but, since we assume that most readers will already have met most of the ideas involved, our treatment will be brisk and will touch only on the essentials.

§1. The language L

The logician De Morgan once gave the following example of an argument, which is undoubtably logically valid, but which cannot be expressed in propositional calculus, that is, in the system described in the previous chapter:

All horses are animals.
Therefore all horses' heads are animals' heads.

In order to cope with valid arguments of this sort involving relations we now introduce the language L of predicate calculus with equality.

The *basic symbols* of L fall into six categories as follows:

(a). *Individual variables.* The individual variables of L are the elements of the countable set

$$\{v_n : n \in \omega\}.$$

Usually we shall call the elements of this set just *variables.*

(b). *Predicate letters.* The predicate letters of L are the elements of the countable set

$$\{P_n : n \in \omega\}.$$

Associated with each predicate letter P_n there is a non-negative integer $\delta(n)$ called the *degree* of P_n.

(c). *Equality symbol.* L has the equality symbol

$$=$$

which is sometimes also called the *identity* symbol.

(d). *Logical connectives.* The logical connectives of L are

$$\neg \text{ (read 'not')} \quad \text{and} \quad \wedge \text{ (read 'and')}$$

(e). *Quantifier symbol.* L has the single quantifier symbol

$$\exists \text{ (read 'there exists')}$$

(*f*). *Punctuation symbols.* The punctuation symbols of L are

$$(\quad) \quad ,$$

By a *logical symbol* we mean a logical connective or a quantifier symbol. A *string* is a finite sequence of symbols of L. We again single out a subset of the strings, the *formulas* of L. We do this in two stages. An *atomic formula* is a string which is either of the form

$$x = y,$$

where x, y are two variables (not necessarily distinct), or of the form

$$P_n(x_1,\ldots,x_{\delta(n)})$$

where $x_1,\ldots,x_{\delta(n)}$ are any $\delta(n)$ variables of L. We are now able to give the following recursive definition of a *formula*:

(1). An atomic formula is a formula.

(2). If ϕ,ψ are formulas and x is a variable, then $\neg\phi$, $(\phi \wedge \psi)$ and $(\exists x)\phi$ are also formulas.

(3). A string is a formula only if it follows from a finite number of applications of rules (1) and (2) that it is a formula.

We let F be the set of formulas of L. Thus F is the smallest set of strings which contains all the atomic formulas and whenever it contains ϕ and ψ also contains $\neg\phi$, $(\phi \wedge \psi)$ and $(\exists x)\,\phi$, where x is any variable.

It should be noted that which strings are formulas depends on the values of the function δ. So we obtain a different language for each function δ defined on the set of natural numbers with natural numbers as values. We could indicate the dependence of the language L on δ by using L_δ to denote this language. However this would cause an unnecessary complication in the notation and so we prefer to be deliberately ambiguous and use just L.

As was the case with propositional calculus we now introduce some new

logical symbols as abbreviatory devices. We introduce the connectives

$$\vee \quad \text{('or')},$$
$$\rightarrow \quad \text{('implies')}$$

and

$$\leftrightarrow \quad \text{('iff')}$$

which are used in the same way as in propositional calculus. That is,

$$(\phi \vee \psi) \text{ abbreviates } \neg(\neg\phi \wedge \neg\psi),$$
$$(\phi \rightarrow \psi) \text{ abbreviates } \neg(\phi \wedge \neg\psi),$$

and

$$(\phi \leftrightarrow \psi) \text{ abbreviates } (\neg(\phi \wedge \neg\psi) \wedge \neg(\psi \wedge \neg\phi)).$$

We also introduce the new quantifier symbol

$$\forall \quad \text{('for all')}$$

which is used in the following way:

$$(\forall x)\, \phi \text{ abbreviates } \neg(\exists x)\neg\phi.$$

Again we omit brackets or use brackets of a different shape when this makes formulas easier to read.

It is clear from these definitions that if ϕ and ψ are formulas and x is any variable then $(\phi \vee \psi)$, $(\phi \rightarrow \psi)$, $(\phi \leftrightarrow \psi)$ and $(\forall x)\, \phi$ are also formulas.

Let x be a variable. By the *x-quantifier* we mean the string

$$(\exists x).$$

We define the *scope* of an occurrence of the x-quantifier in a formula recursively as follows. In each case we underline the occurrence of the x-quantifier whose scope is being defined:

(a). The scope of $(\exists x)$ in $\underline{(\exists x)}\phi$ is $(\exists x)\phi$.

(b). The scope of $(\exists x)$ in $\neg X\,\underline{(\exists x)}\,Y$ is the same as its scope in $X\,\underline{(\exists x)}\,Y$.

(c). The scope of $(\exists x)$ in $X\underline{(\exists x)}\,Y \wedge \psi$ or $\psi \wedge X\underline{(\exists x)}\,Y$ is the same as its scope in $X\underline{(\exists x)}\,Y$.

(d). The scope of $(\exists x)$ in $(\exists y)X\underline{(\exists x)}\,Y$ is the same as its scope in $X\underline{(\exists x)}\,Y$.

Here X and Y are any (possibly empty) strings, ϕ and ψ are any formulas and y is any variable. Thus, for example, in the formula

$$\neg(\exists v_0)[(\exists v_1)[(\exists v_2)\,P(v_1,v_2) \wedge P(v_1,v_0)] \vee \neg P(v_0,v_0)]$$

the scope of the v_0-quantifier is

$$(\exists v_0)[(\exists v_1)[(\exists v_2)\,P(v_1,v_2) \wedge P(v_1,v_0)] \vee \neg P(v_0,v_0)],$$

that of the v_1-quantifier is

$$(\exists v_1)[(\exists v_2)P(v_1,v_2) \wedge P(v_1,v_0)],$$

and that of the v_2-quantifier is

$$(\exists v_2)P(v_1,v_2).$$

An occurrence of the variable x in a formula is said to be a *bound* occurrence if it occurs in the scope of an x-quantifier. Otherwise it is said to be a *free* occurrence of the variable. Thus given any formula ϕ and any occurrence of the variable x in ϕ, by seeing how ϕ is built up from atomic formulas we can tell whether the occurrence of x is bound or free. Any particular occurrence is either free or bound, but not both, however a variable x can have both free and bound occurrences in the same formula.

EXERCISES

1.1. In the formula

$$(\exists v_1)[(\forall v_2)[v_1 = v_2 \rightarrow P_1(v_2,v_3)] \vee [v_1 = v_4 \wedge P_2(v_2,v_4)]],$$

which occurrences of variables are free and which are bound?

1.2. Show that the following definition of free occurrences of variables is equivalent to that given above:

(a). All occurrences of variables in atomic formulas are free occurrences.

(b). The free occurrences of x in $\neg\phi$ correspond to the free occurrences of x in ϕ.

(c). The free occurrences of x in $\phi \wedge \psi$ correspond to the free occurrences of x in ϕ and the free occurrences of x in ψ.

(d). The free occurrences of x in $(\exists y)\phi$ correspond to the free occurrences of x in ϕ if y is a different variable to x, while if y is the same as x then x has no free occurrences in $(\exists y)\phi$.

Suppose that ϕ is a formula which has no free occurrence of the variable x. Then we say that the first occurrence of the x-quantifier in the formula $(\exists x)\phi$ is a *vacuous* occurrence of that quantifier. It should be noted that our formation rules allow vacuous occurrences of quantifiers. Thus according to our rules both

$$(\exists v_0)P_1(v_1,v_2) \quad \text{and} \quad (\exists v_0)(\exists v_0)P_1(v_0,v_1)$$

are (well-formed) formulas (provided, of course, that P_1 is a predicate letter of degree 2).

A formula in which no variable occurs free is called a *sentence*. If the

formula ϕ contains the variable x free, then, in a sense which will be made more precise later on, ϕ 'says something' about x and we often indicate this situation by writing ϕ as $\phi(x)$. We denote the result of replacing all free occurrences of x in $\phi(x)$ by occurrences of the variable y by $\phi(y)$. $\phi(y)$ will not necessarily say the same thing about y as $\phi(x)$ says about x. Compare, for example, the following pairs of formulas:

$\phi(v_0)$	$\phi(v_1)$
$P_1(v_0,v_2)$	$P_1(v_1,v_2)$
$P_2(v_0,v_1)$	$P_2(v_1,v_1)$
$(\exists v_1)[P_3(v_0)\rightarrow P_4(v_1)]$	$(\exists v_1)[P_3(v_1)\rightarrow P_4(v_1)]$

Only in the first case does $\phi(v_1)$ say the same thing about v_1 as $\phi(v_0)$ does about v_0. We see from these examples that for this to be the case we must have that

(a). v_1 does not occur free in $\phi(v_0)$, and

(b). v_0 does not have any free occurrences in $\phi(v_0)$ which are in the scope of a v_1-quantifier.

If (b) holds we say that v_1 is *free for* v_0 in $\phi(v_0)$. It is evident that, in general, y is free for x in ϕ if whenever all free occurrences of x in ϕ are replaced by occurrences of y then the new occurrences of y that result from this substitution are all free occurrences.

EXERCISE 1.3. In which of the following formulas is v_1 free for v_0?

(a). $(\exists v_2)P(v_0,v_2)$;

(b). $(\forall v_0)(\exists v_2)P(v_0,v_2)$;

(c). $(\exists v_1)(\exists v_0)P(v_0,v_1)\wedge P(v_1,v_0)$;

(d). $(\forall v_1)[(\exists v_0)P(v_0,v_1)\wedge P(v_1,v_0)]$;

(e). $(\exists v_1)P(v_1,v_1)$.

§2. Interpretations of L

We saw in the previous chapter that, essentially, interpretations of formulas of propositional calculus consist of assignments of the truth values *true* and *false* to the propositional variables occurring in them. In the case of predicate calculus the situation is more complicated. In order to interpret formulas of L we have to specify firstly over which set of objects we take the variables

to be ranging and secondly how the predicate letters are to be interpreted. Both choices will affect the truth of a formula under a given interpretation. Consider, for example, the sentence σ, namely

$$(\forall v_0)\, M(v_0).$$

If we interpret M as the predicate '... is a European', then σ is true if we interpret the variables as ranging over all Englishmen and false if we interpret the variables as ranging over all algebraists. While if we interpret the variables as ranging over all algebraists then σ is false if we interpret M as the predicate '... is an Englishman' and true if we interpret M as '... is a mathematician'. Thus an interpretation of L will consist of a set over which we consider the variables to range, and predicates or relations defined on this set as interpretations of the predicate letters.

By a *relational structure* we shall mean such an interpretation. That is, a *relational structure* is an ordered pair

$$\mathfrak{A} = \langle A, \{R_n : n \in \omega\} \rangle,$$

where A, the *domain* of $\mathfrak{A}$, is a *non-empty* set and for $n \in \omega$ R_n is a finitary, say $\lambda(n)$-ary, relation on A. We reserve upper case German letters to denote relational structures and we use the corresponding Roman letter, with the necessary subscripts, to denote the domain of the given relational structure. The relational structure $\mathfrak{A}$ will count as an interpretation of the language L if the degrees of the relations R_n correspond to the degrees of the predicate letters P_n. That is, if for $n \in \omega$, $\delta(n) = \lambda(n)$. In this case we say that the relational structure $\mathfrak{A}$ is a *realization* of the language L and that L is *appropriate* for the structure $\mathfrak{A}$. We call the relation R_n the *value* of P_n in the realization $\mathfrak{A}$.

If one was asked whether

$$x < 9$$

was a true formula of arithmetic the natural response would be: 'it all depends on the value of x'. We see from this that if $\phi(v)$ is a formula containing a free occurrence of the variable v then before we can say whether or not this formula is 'true' of the realization $\mathfrak{A}$ we have to interpret v as referring to some specific element of A. It is most convenient to do this by giving a simultaneous interpretation to all the variables of L whenever they occur free in formulas. This is done by specifying a sequence

$$x = \langle x_0, x_1, \ldots, x_n, \ldots \rangle$$

of elements of A. x is called a *valuation* of the variables of L. The valuation

x tells us to interpret free occurrences of the variable v_n as referring to the element x_n of A. If x is a valuation, and $a \in A$, $x(n/a)$ is the valuation which assigns the same values to the variables as does x except that it assigns the value a to v_n. Thus if x is as above, then

$$x(n/a) = \langle x_0, x_1, \ldots, x_{n-1}, a, x_{n+1}, \ldots \rangle .$$

Notice that the value assigned to v_n by $x(n/a)$ is independent of the value assigned to v_n by x.

We are now able to give a precise definition, due to Tarski, of whether a formula is true of a given realization of L under a given valuation of the free variables. We define the relation x *satisfies* ϕ in $\mathfrak{A}$, which we write as $\mathfrak{A} \vDash_x \phi$, recursively as follows:

(a). $\mathfrak{A} \vDash_x v_m = v_n$ iff x_m is the same element as x_n.
(b). $\mathfrak{A} \vDash_x P_n(v_{i_1}, \ldots, v_{i_{\delta(n)}})$ iff $\langle x_{i_1}, \ldots, x_{i_{\delta(n)}} \rangle \in R_n$.
(c). $\mathfrak{A} \vDash_x \neg \phi$ iff not $\mathfrak{A} \vDash_x \phi$.
(d). $\mathfrak{A} \vDash_x \phi \wedge \psi$ iff $\mathfrak{A} \vDash_x \phi$ and $\mathfrak{A} \vDash_x \psi$.
(e). $\mathfrak{A} \vDash_x (\exists v_n)\phi$ iff for some $a \in A$, $\mathfrak{A} \vDash_{x(n/a)} \phi$.

EXERCISES

2.1. Let ϕ be a formula of L, $\mathfrak{A}$ be a realization of L and x a valuation. Show that $\mathfrak{A} \vDash_x (\forall v_n)\phi$ iff for all $a \in A$, $\mathfrak{A} \vDash_{x(n/a)} \phi$.

2.2. Let ϕ be a formula of L, $\mathfrak{A}$ be a realization of L, and let x, y be valuations such that whenever v_n occurs free in ϕ, $x_n = y_n$. Prove that $\mathfrak{A} \vDash_x \phi$ iff $\mathfrak{A} \vDash_y \phi$.

Because of the result given by this last exercise we shall often write $\mathfrak{A} \vDash \phi[x_0, \ldots, x_n]$ instead of $\mathfrak{A} \vDash_x \phi$ whenever the free variables of ϕ occur among $v_0, \ldots, v_n$. Thus we write $\mathfrak{A} \vDash \phi[x_0, \ldots, x_n]$ whenever any valuation which assigns v_i the value x_i, for $i \leqslant n$, satisfies ϕ in $\mathfrak{A}$. In particular if σ is a sentence, that is a formula without any free variables that require interpretation, then if $\mathfrak{A} \vDash_x \sigma$ for *some* valuation x, then $\mathfrak{A} \vDash_x \sigma$ for *all* valuations x. In this case we just write $\mathfrak{A} \vDash \sigma$ and we say that σ is *true* or *valid* in $\mathfrak{A}$ and that $\mathfrak{A}$ is a *model* of σ. If Σ is a set of sentences of L we say that a realization $\mathfrak{A}$ is a *model* of Σ if $\mathfrak{A}$ is a model of each sentence in Σ. In this case we write $\mathfrak{A} \vDash \Sigma$.

A sentence σ is said to be *universally valid* if it is valid in all realizations. Thus, roughly speaking, a sentence is universally valid if it is true under any interpretation. We shall want to qualify this last remark slightly later on. The notion of a universally valid sentence of predicate calculus is therefore

analogous to that of a tautology of propositional calculus. In the case of propositional calculus we saw that in order to determine whether or not a given formula is a tautology we need only check a finite number of realizations and thus that we had a decision procedure for determining whether or not a formula is a tautology. The situation is different for predicate calculus.

Although we show subsequently that in order to decide whether a sentence of predicate calculus is universally valid it is not necessary to check all realizations, it will, in general, still be necessary to check an infinite number of them. Also, given any particular realization there is not necessarily a decision procedure for deciding whether an arbitrary sentence is true in it. Therefore the method of checking realizations does not provide a decision procedure for universal validity. Indeed it has been shown by CHURCH [1936] that there is no decision procedure for determining whether or not an arbitrary sentence of predicate calculus is universally valid. In other words, the set of universally valid sentences of predicate calculus is not recursive. We do not give a proof of this result since it is outside the scope of the book.

Given this situation there is considerable interest in trying to find an axiom system that will generate all universally valid sentences. Among other things, if we succeed in this task we shall have shown that the universally valid sentences of predicate calculus form a recursively enumerable set, a fact which is not evident from the definition. We turn our attention to this task in the next section.

§ 3. An axiom system for predicate calculus

We now describe the axiom system P C for predicate calculus with equality.

The axioms of PC fall into three groups. By a *substitution instance of a propositional tautology* we mean a formula of L which can be obtained from a tautology of propositional calculus by substituting formulas of L for the propositional variables. It is clear that all substitution instances of propositional tautologies are universally valid. To ensure that they are all provable in PC we take as our first group of axioms all instances of the following schemas, corresponding to the axiom schemas of SC:

PC1. $\phi \to (\psi \to \phi)$,

PC2. $[\phi \to (\psi \to \chi)] \to [(\phi \to \psi) \to (\phi \to \chi)]$,

PC3. $(\neg\phi \to \neg\psi) \to (\psi \to \phi)$,

PC4a. $(\phi \wedge \psi) \to \phi$,

PC4b. $(\phi \wedge \psi) \to \psi$,

PC5. $(\chi\to\phi)\to[(\chi\to\psi)\to(\chi\to[\phi\wedge\psi])]$,
PC6a. $\phi\to(\phi\vee\psi)$,
PC6b. $\psi\to(\phi\vee\psi)$,
PC7. $(\phi\to\chi)\to[(\psi\to\chi)\to([\phi\vee\psi]\to\chi)]$,
where ϕ,ψ and χ are any formulas of L.

Secondly we require some axioms involving the quantifiers of L. We take for these all instances of the two schemas:
PC8. $(\forall x)\phi(x)\to\phi(y)$,
PC9. $(\forall x)(\phi\to\psi)\to(\phi\to(\forall x)\psi)$,
where x,y are any variables of L, ψ is any formula of L and in PC8 ϕ is any formula of L in which y is free for x and in PC9 ϕ is any formula of L which contains no free occurrence of the variable x.

Finally we require axioms involving the equality symbol. We take for these all instances of the following schemas:
PC10. $(\forall x)(x=x)$,
PC11. $(\forall x)(\forall y)(x=y\to[\phi\to\phi'])$,
where x is any variable, ϕ is any formula of L in which the variable y is free for x, and ϕ' is the result of replacing some, but not necessarily all, of the free occurrences of x in ϕ by y.

The system PC has two rules of inference. These are:

Modus ponens: ψ is an *immediate consequence* of ϕ and $\phi\to\psi$, where ϕ and ψ are any formulas of L.

Generalization: $(\forall x)\phi$ is an *immediate consequence* of ϕ, where ϕ is any formula and x is any variable of L.

A *proof* of a formula ϕ is a finite sequence, say $\phi_1,\ldots,\phi_n$, of formulas such that $\phi=\phi_n$ and for each $i\leqslant n$, ϕ_i either is an axiom, or for some $j,k<i$ is an immediate consequence of ϕ_j and ϕ_k according to the rule modus ponens, or for some $j<i$ is an immediate consequence of ϕ_j according to the rule of generalization. A formula ϕ is said to be *provable* if there is a proof of it. If ϕ is a provable formula we write $\vdash\phi$.

LEMMA 3.1. *If ϕ is a substitution instance of a propositional tautology then* $\vdash\phi$.

PROOF. Since all formulas of the forms PC1–PC7 are axioms of PC and modus ponens is a rule of inference of PC, this is obvious.

We use this result frequently to simplify proofs that formulas are provable in PC. We have already adopted the convention of writing $\phi(x)$ to indicate that the variable x occurs free in ϕ. We extend this convention and write $\phi(x_1,\ldots,x_n)$ to indicate that all the variables $x_1,\ldots,x_n$ occur free in ϕ.

LEMMA 3.2. *For any formula* $\phi(x, y)$ *of* L,

$$\vdash(\forall x)(\forall y)\,\phi(x,y)\rightarrow(\forall y)(\forall x)\,\phi(x,y).$$

PROOF. By PC8,

$$\vdash(\forall x)(\forall y)\,\phi(x,y)\rightarrow(\forall y)\,\phi(x,y),$$

and

$$\vdash(\forall y)\,\phi(x,y)\rightarrow\phi(x,y).$$

Therefore, using lemma 3.1, the fact that $(p\rightarrow q)\rightarrow[(q\rightarrow r)\rightarrow(p\rightarrow r)]$ is a propositional tautology, and two applications of modus ponens,

$$\vdash(\forall x)(\forall y)\,\phi(x,y)\rightarrow\phi(x,y).$$

Hence, by generalization

$$\vdash(\forall x)[(\forall x)(\forall y)\,\phi(x,y)\rightarrow\phi(x,y)],$$

and so, by PC9 and modus ponens,

$$\vdash(\forall x)(\forall y)\,\phi(x,y)\rightarrow(\forall x)\,\phi(x,y).$$

Finally, using generalization, PC9, and modus ponens again in a similar way

$$\vdash(\forall x)(\forall y)\,\phi(x,y)\rightarrow(\forall y)(\forall x)\,\phi(x,y).$$

By an *open* formula we shall mean a formula which contains no quantifiers. The *closure* of a formula ϕ is the formula which results from ϕ on prefixing the universal quantifier $(\forall x)$ for each variable x that occurs free in ϕ. Thus if the free variables that occur in ϕ are $x_1,\ldots,x_n$ then the closure of ϕ is

$$(\forall x_1)\ldots(\forall x_n)\,\phi.$$

By lemma 3.2 the order in which we prefix the quantifiers is unimportant. The closure of a formula contains no free variables and is therefore a sentence.

LEMMA 3.3. *A formula is provable iff its closure is provable.*

PROOF. This is evident since PC contains the rule of generalization and the axiom schema PC8.

Let Σ be a set of formulas of L. A formula ϕ is said to be *provable from the assumptions* Σ if there is a finite sequence of formulas, say $\phi_1,\ldots,\phi_n$, such that $\phi=\phi_n$ and for each $i\leqslant n$, ϕ_i either is an axiom, one of the formulas in Σ, or for some $j,k<i$ is an immediate consequence of ϕ_j and ϕ_k according to the rule modus ponens, or for some $j<i$ is an immediate consequence of

ϕ_j according to the rule generalization. If ϕ is provable from the assumptions Σ we write $\Sigma \vdash \phi$, and we say that ϕ is *deducible* from Σ.

THEOREM 3.4. THE FINITENESS THEOREM FOR PREDICATE CALCULUS
If $\Sigma \vdash \phi$, there is some finite subset Σ_0 of Σ such that $\Sigma_0 \vdash \phi$.

PROOF. The proof of this result is the same as that of the finiteness theorem for propositional calculus.

EXERCISES

3.5. Show that for any variables x,y,z of L

$$\vdash (\forall x)(\forall y)[x=y \to y=x],$$

and

$$\vdash (\forall x)(\forall y)(\forall z)[(x=y \wedge y=z) \to x=z].$$

3.6. Show that if $\phi(x_1,\ldots,x_n)$ is any formula with free variables among $x_1,\ldots,x_n$ and for $i \leqslant n$, y_i is free for x_i in ϕ then

$$\vdash [x_1=y_1 \wedge \cdots \wedge x_n=y_n] \to [\phi(x_1,\ldots,x_n) \to \phi(y_1,\ldots,y_n)].$$

We let *Val* be the set of all universally valid sentences of L and *Prov* be the set of all provable sentences of L. We now begin investigating the relationship between these two sets.

THEOREM 3.7. *Each provable sentence is universally valid.*

PROOF. The proof is essentially the same as that of the analogous theorem for propositional calculus (theorem 2.1.11) by induction on the length of proofs. However not all the formulas occurring in a proof of a sentence will necessarily be sentences. So what we in fact show is that the closure of each formula occurring in a proof is universally valid. To prove this it is sufficient to check that the closure of any axiom is universally valid and that the rules of inference yield formulas whose closures are universally valid from formulas which have this property. This we leave as an exercise for the reader.

A set Σ of formulas is said to be *consistent* if there is no formula ϕ such that both ϕ and $\neg\phi$ are provable from the assumptions Σ.

COROLLARY 3.8. THE CONSISTENCY OF PREDICATE CALCULUS
The empty set of formulas is consistent.

PROOF. By theorem 3.7 if both ϕ and $\neg\phi$ are provable without any assumptions then both ϕ and $\neg\phi$ are universally valid which is clearly impossible.

This corollary shows that the system PC is consistent since it cannot be used to prove contradictory formulas. The proof of consistency is not constructive since it involves the notion of universal validity. Readers who desire a constructive consistency proof should observe that our proof remains correct if we replace 'universally valid' by 'valid in all realizations whose domain contains just one element' and then that it can be effectively checked whether or not a given sentence is valid in all such realizations.

EXERCISE 3.9. Suppose that $\Sigma \vdash \phi$ and $\mathfrak{A}$ is a realization of L. If the closure of each formula in Σ is true in $\mathfrak{A}$, show that the closure of ϕ is also true in $\mathfrak{A}$.

Theorem 3.7 shows that *Prov* $\subseteq$ *Val*. We devote the next two sections to showing that the converse inclusion also holds.

§4. The Lindenbaum algebra of PC

We now construct the Lindenbaum algebra $\mathscr{A}$ of PC in just the same way that we used for propositional calculus in the previous chapter. Recall that F is the set of all formulas of L. We define the relation $\equiv$ on F by

$$\phi \equiv \psi \quad \text{iff} \quad \vdash \phi \to \psi \quad \text{and} \quad \vdash \psi \to \phi .$$

$\equiv$ is an equivalence relation on F. If ϕ is a formula we let $|\phi|$ be the equivalence class under the relation $\equiv$ to which ϕ belongs, and we let $F/\equiv$ be the set of equivalence classes of the formulas in F. Next we define the relation $\leqslant$ on $F/\equiv$ by

$$|\phi| \leqslant |\psi| \quad \text{iff} \quad \vdash \phi \to \psi .$$

As was the case with propositional calculus this is an unambiguous definition and $\mathscr{A} = \langle F/\equiv, \leqslant \rangle$ is a Boolean algebra, the *Lindenbaum algebra* of PC.

For each formula ϕ in F let $\phi(v_{k_1}/v_{p_1},\ldots,v_{k_n}/v_{p_n})$ be the formula obtained from ϕ by (1) replacing all bound occurrences of v_{p_i} by v_{j+i}, where j is the least number exceeding $p_1,\ldots,p_n$ and the subscripts of all the variables occurring in ϕ, and then (2) replacing all free occurrences of v_{k_i} by v_{p_i}, for $1 \leqslant i \leqslant n$.

LEMMA 4.1. *For each formula* ϕ *of* L,

$$|(\forall v_k)\phi| = \inf\{|\phi(v_k/v_p)| : p \in \omega\} .$$

PROOF. By PC8,

$$\vdash (\forall v_k)\phi \to \phi(v_k/v_p),$$

and so, for $p \in \omega$,

$$|(\forall v_k)\phi| \leqslant |\phi(v_k/v_p)|,$$

whence

$$|(\forall v_k)\phi| \leqslant \inf\{|\phi(v_k/v_p)| : p \in \omega\}.$$

Now suppose that ψ is a formula such that $|\psi|$ is a lower bound of the set $\{|\phi(v_k/v_p)| : p \in \omega\}$. Then, in particular, choosing q so that v_q does not occur in ϕ or ψ,

$$\vdash \psi \to \phi(v_k/v_q),$$

whence, by generalization and PC9,

$$\vdash \psi \to (\forall v_q)\phi(v_k/v_q).$$

By PC8

$$\vdash (\forall v_q)\phi(v_k/v_q) \to \phi,$$

whence, by generalization and PC9,

$$\vdash (\forall v_q)\phi(v_k/v_q) \to (\forall v_k)\phi.$$

It follows that

$$\vdash \psi \to (\forall v_k)\phi,$$

and thus $|(\forall v_k)\phi|$ is the greatest lower bound, i.e. the infimum, of the set $\{|\phi(v_k/v_p)| : p \in \omega\}$, as was to have been proved.

The importance of this lemma will become apparent in the next section.

EXERCISE 4.2. Let D be an ultrafilter in the Lindenbaum algebra $\mathscr{A}$. Show that for any formulas ϕ, ψ of L:

(a). $|\neg\phi| \in D$ iff $|\phi| \notin D$,

(b). $|\phi \wedge \psi| \in D$ iff $|\phi| \in D$ and $|\psi| \in D$,

and

(c). if $|\phi|, |\phi \to \psi| \in D$ then $|\psi| \in D$.

§ 5. The completeness of PC.

We have seen that *Prov* $\subseteq$ *Val*. We now show that our system PC is complete by proving that the converse inclusion also holds.

THEOREM 5.1. THE COMPLETENESS THEOREM FOR PREDICATE CALCULUS
Each universally valid sentence is provable.

PROOF. Let σ be an unprovable sentence of L. We show that there is a relational structure in which σ is not true. This will show that σ is not universally valid.

Since not $\vdash \sigma$, $|\sigma| \neq 1$ in the Lindenbaum algebra $\mathscr{A}$ and so $|\neg\sigma| \neq 0$ in this algebra. By lemma 4.1, for each formula ϕ of L,

$$|(\forall v_k)\phi| = \inf\{|\phi(v_k/v_p)| : p \in \omega\} \qquad [\Pi]$$

There are only countably many formulas of L. It therefore follows from Tarski's lemma that there is an ultrafilter D in $\mathscr{A}$ which contains $|\neg\sigma|$ and preserves all the meets $[\Pi]$. Note that this implies that

$$|(\forall v_k)\phi| \in D \quad \text{iff} \quad \text{for all } p \in \omega,\ |\phi(v_k/v_p)| \in D.$$

We define the realization $\mathfrak{A}$ of L as follows:

Let V be the set of variables of L. We define the relation $\sim$ on V by

$$v_i \sim v_j \quad \text{iff} \quad |v_i = v_j| \in D.$$

By exercises 3.5 and 4.2, $\sim$ is an equivalence relation on V.

For $v \in V$ let

$$v^{\sim} = \{v' \in V : v' \sim v\},$$

and let

$$V^{\sim} = \{v^{\sim} : v \in V\}.$$

We let $V^{\sim}$ be the domain of $\mathfrak{A}$. Clearly $V^{\sim}$ is non-empty and is of cardinal $\leqslant \aleph_0$.

For each $n \in \omega$ we define the relation R_n on V by

$$\langle w_1{}^{\sim}, \ldots, w_{\delta(n)}{}^{\sim} \rangle \in R_n \quad \text{iff} \quad |P_n(w_1, \ldots, w_{\delta(n)})| \in D.$$

By exercise 3.6 this is an unambiguous definition, that is, if for $1 \leqslant i \leqslant \delta(n)$, $u_i \sim w_i$, then

$$|P_n(u_1, \ldots, u_{\delta(n)})| \in D \quad \text{iff} \quad |P_n(w_1, \ldots, w_{\delta(n)})| \in D.$$

Then we put

$$\mathfrak{A} = \langle V^{\sim}, \{R_n : n \in \omega\} \rangle.$$

If $x \in V^{\omega}$, say $x = \langle x_0, \ldots, x_n, \ldots \rangle$, then we let $x^{\sim} \in (V^{\sim})^{\omega}$ be the sequence $\langle x_0{}^{\sim}, \ldots, x_n{}^{\sim}, \ldots \rangle$ of elements of $V^{\sim}$. Clearly all sequences of elements of $V^{\sim}$ are of this form.

We now show that for each formula $\phi(v_0, \ldots, v_n)$ of L, whose free variables are among $v_0, \ldots, v_n$,

$$\text{for all } x \in V^{\omega}, \quad \mathfrak{A} \vDash_{x^{\sim}} \phi \quad \text{iff} \quad |\phi(v_0/x_0, \ldots, v_n/x_n)| \in D \qquad [*]$$

The proof of [*] will be by induction on the number of logical symbols in ϕ. We consider first the case of atomic formulas.

If ϕ is an atomic formula of the form $v_i = v_j$ then it follows from the definition of $V^\sim$ that [*] holds, while if ϕ is of the form $P_n(w_1,\ldots,w_{\delta(n)})$ this follows from the definition of the relation R_n.

Now suppose that [*] holds for all formulas containing fewer than μ logical symbols and that ϕ is a formula containing precisely μ logical symbols. There are three cases to consider: ϕ is of the form $\neg\psi$, $\psi \wedge \chi$ or $(\exists v_n)\psi$, where ψ and χ are formulas containing fewer than μ logical symbols and are thus, by hypothesis, formulas for which [*] holds. If ϕ is of the form $\neg\psi$ or $\psi \wedge \chi$, then it follows at once from exercise 4.2 that [*] holds also for ϕ.

Assume now that ϕ is $(\exists v_n)\psi$, where ψ is a formula whose free variables are among $v_0,\ldots,v_n$. Then

$$\begin{aligned} \mathfrak{A} \vDash_{x^\sim} \phi \quad &\text{iff} \quad \text{for some } v^\sim \in V^\sim,\ \mathfrak{A} \vDash_{x^\sim (n/v^\sim)} \psi \\ &\text{iff} \quad \text{for some } p \in \omega,\ \mathfrak{A} \vDash_{x^\sim (n/v_p^\sim)} \psi\,, \end{aligned}$$

and by our induction hypothesis this holds iff

$$\text{for some } p \in \omega, \quad |\psi(v_0/x_0,\ldots,v_{n-1}/x_{n-1},v_n/v_p)| \in D\,. \tag{1}$$

Now (1) holds iff

$$\text{not for all } p \in \omega, \quad |\neg\psi(v_0/x_0,\ldots,v_{n-1}/x_{n-1},v_n/v_p)| \in D\,, \tag{2}$$

and since D preserves all the meets $[\Pi]$, (2) holds iff

$$\text{not} \quad |(\forall v_n)\neg\psi(v_0/x_0,\ldots,v_{n-1}/x_{n-1},v_n)| \in D\,,$$

that is, iff

$$|\neg(\forall v_n)\neg\psi(v_0/x_0,\ldots,v_{n-1}/x_{n-1},v_n)| \in D\,.$$

Therefore, since

$$|(\exists v_n)\psi\,(v_0/x_0,\ldots,v_{n-1}/x_{n-1},v_n)| = |\neg(\forall v_n)\neg\psi\,(v_0/x_0,\ldots,v_{n-1}/x_{n-1},v_n)|\,,$$

we have shown that [*] holds also for ϕ.

We can now deduce that [*] holds for all formulas. In particular, since $|\neg\sigma| \in D$, $\mathfrak{A}$ is a model of $\neg\sigma$. Therefore σ is not universally valid and our proof of the theorem is complete.

A sentence σ is said to be *refutable* if $\neg\sigma$ is provable and *irrefutable* if $\neg\sigma$ is not provable. A relational structure is said to be *of cardinal* α if its domain is of cardinal α, and is said to be *countable* if its domain is countable.

COROLLARY 5.2. *A sentence of* L *is irrefutable iff it has a countable model.*

PROOF. If σ is refutable $\neg\sigma$ is universally valid and so σ has no models at all.

If σ is irrefutable then as in the proof of the completeness theorem we can construct a model of σ. Notice that since V is countable, $V^{\sim}$ is also countable and thus the model $\mathfrak{A}$ of σ that we construct is countable ($\mathfrak{A}$ may be finite).

COROLLARY 5.3. *A finite set of sentences of* L *is consistent iff it has a countable model.*

PROOF. If $\Sigma = \{\sigma_1,\ldots,\sigma_n\}$ is a finite set of sentences of L it is clear from the results of the previous chapter that Σ is consistent if the sentence $\sigma_1 \wedge \ldots \wedge \sigma_n$ is irrefutable. The result is now an immediate consequence of the previous corollary.

Remark (which presupposes some knowledge of axiomatic set theory). The completeness theorem together with theorem 3.7 implies that

$$Val = Prov.$$

Recall that *Val* is the set of universally valid sentences of L, that is the set of sentences true in all relational structures. The domain of a relational structure is a *set*. However the original intention when predicate calculus was first formulated was to capture the notion of a *logically valid* sentence, that is, a sentence true under *any* interpretation. For example, a logically valid sentence must be true in the universe of all sets (when its predicate letters are given interpretations) and the domain of this interpretation is not a set (e.g. in Gödel–Bernays set theory) and so this interpretation is not a relational structure according to our definition. Thus logical validity is a stricter notion than universal validity, and if V is the set of logically valid sentences then *a priori* all we have is

$$Prov \subseteq V \subseteq Val.$$

However, now that we have proved that $Prov = Val$ we can conclude that

$$Prov = V = Val.$$

and so the system PC is indeed sufficient to prove all logically valid sentences. This considerably increases the interest of the completeness theorem. Compare, for example, the situation with modal logic. Here various elegant completeness results have recently been obtained but it is still a matter of controversy as to whether any particular modal system adequately captures

the notion of 'necessity'. But the remarks above show quite conclusively that PC captures the notion of logical validity for sentences of the language L.

Now let Val_0 be the set of sentences of L which are true in all countable relational structures.

COROLLARY 5.4. $Val_0 = Val = V = Prov$.

PROOF. It is sufficient to prove $Val_0 = Prov$. Clearly $Prov \subseteq Val_0$. To see that the converse is true notice that if $\sigma \notin Prov$ then $\neg\sigma$ is irrefutable and so has a countable model. This implies that $\sigma \notin Val_0$.

This surprising result shows that validity for sentences of first order predicate calculus under *any* interpretation can be reduced to validity in *countable* structures. It is clear that *countable* cannot be replaced by *finite* here. In ch. 13 we give an example of a predicate language for which logical validity cannot be reduced to validity in countable structures.

The problem of generalizing corollary 5.3 to infinite sets of sentences is much more difficult than was the analogous problem for propositional calculus and we postpone its solution until ch. 5. The deduction theorem on the other hand can be extended to predicate calculus in a straightforward way.

THEOREM 5.5. THE DEDUCTION THEOREM FOR PREDICATE CALCULUS
Let Σ be a set of sentences of L *and let σ,τ be sentences of* L *such that*

$$\Sigma \cup \{\sigma\} \vdash \tau,$$

then

$$\Sigma \vdash \sigma \to \tau.$$

PROOF. By the finiteness theorem there is some finite subset $\Sigma_0 = \{\sigma_1,\ldots,\sigma_n\}$ of Σ such that $\Sigma_0 \cup \{\sigma\} \vdash \tau$. By exercise 3.9 any model of $\Sigma_0 \cup \{\sigma\}$ also satisfies τ. It follows that $(\sigma_1 \wedge \ldots \wedge \sigma_n) \to (\sigma \to \tau)$ is universally valid and so, by the completeness theorem, provable. It follows at once that $\Sigma_0 \vdash \sigma \to \tau$ and so $\Sigma \vdash \sigma \to \tau$.

It should be noted that this result does not hold in general for *formulas* of L. For example, by generalization

$$\phi(v_0) \vdash (\forall v_0)\phi(v_0),$$

but clearly the closure of $\phi(v_0) \to (\forall v_0)\phi(v_0)$ is not universally valid and so $\phi(v_0) \to (\forall v_0)\phi(v_0)$ is not provable. It is possible to extend theorem 5.5

to formulas by putting restrictions on the nature of the proof of τ from $\Sigma \cup \{\sigma\}$. Roughly speaking, we would have to insist that generalization is not used with respect to variables that have free occurrences in σ. We shall not need to use this extension of the deduction theorem here. The details can be found in any standard text on predicate calculus, e.g. MENDELSON [1964] page 61.

§ 6. Uncountable languages

In two respects the language L is more powerful than other first order predicate languages we might have considered. First, L has the equality symbol $=$ which is always interpreted as the identity relation. It is possible to consider predicate languages without equality but for our purposes they are not of interest. Second, L has an infinite number of predicate letters and so its realizations are non-empty sets together with an infinite number of relations defined on the set. Most mathematical structures, however, consist of sets with just a finite number of relations. To talk about these structures we only require languages with a finite number of predicate letters. We shall use such languages in the sequel. It is clear that all the results of the previous sections apply to these languages.

There are also several respects in which we can consider more general first order predicate languages. We consider some of these in this and the next two sections.

We have formulated L so that it contains only countably many distinct predicate letters. We used this fact essentially in the proof of the completeness theorem when we said that L has only countably many different formulas and so there are only countably many meets [Π].

Suppose now that we have set up the language L* just like L except that L* has an arbitrary infinite set $\{P_j : j \in J\}$ of predicate letters. We show that the completeness theorem still holds for this language.

The chief thing to notice is that whether or not a sentence σ is true in a relational structure $\mathfrak{A}$ depends only on the values of the predicate letters occurring in σ and not on those of the other predicate letters. So suppose σ is an irrefutable sentence of L*. Let L′ be a countable language containing all predicate letters which occur in σ (necessarily there are only a finite number of them). σ is an irrefutable sentence of L′ and so, by the completeness theorem, there is a countable realization, say $\mathfrak{A}'$, of L′ in which σ is true. We extend $\mathfrak{A}'$ to a realization of L* by defining the values of the predicate letters of L* not in L′ in any way we like (for definiteness let them

all be the empty relation). By our remarks above this realization of L* is still a model of σ. Thus the completeness theorem extends to L*.

§7. Constants

For many purposes it is useful to be able to single out individual elements of relational structures. To enable us to do this we add to our language L a sequence

$$\langle c_\xi : \xi < \beta \rangle$$

of *constant symbols*, also called just *constants*. We denote this extended language by L_β.

Since we have extended the set of basic symbols we need also to extend our formation rules and the axioms. For L_β an *atomic formula* is a string of the form

$$t_1 = t_2 ,$$

or of the form

$$P_n(t_1,\ldots,t_{\delta(n)}),$$

where now $t_1,\ldots,t_{\delta(n)}$ can be either variables or constants. The notion of a *formula* is defined recursively in terms of atomic formulas just as for L. The axioms for the extended language are the same as those for L, except that now in PC8 y can be either a variable or a constant. There is no change in the rules of inference.

A *realization* of L_β is a triple

$$\mathfrak{A}_\beta = \langle A, \{R_n : n \in \omega\}, \langle a_\xi : \xi < \beta \rangle \rangle ,$$

where

$$\mathfrak{A} = \langle A, \{R_n : n \in \omega\} \rangle$$

is a realization of L and

$$a = \langle a_\xi : \xi < \beta \rangle$$

is a β-termed sequence of elements of A, so $a \in A^\beta$. We often use $(\mathfrak{A},a)$ to denote the realization of L_β which is obtained from the realization $\mathfrak{A}$ of L by adding the sequence a of elements of A.

The notion of *satisfaction* is extended to an atomic formula of L_β in the following way:

$(\mathfrak{A},a) \models_x t_i = t_j$ iff b_i is the same element as b_j,

$(\mathfrak{A},a) \models_x P_n(t_1,\ldots,t_{\delta(n)})$ iff $\langle b_1,\ldots,b_{\delta(n)} \rangle \in R_n$,

where if t_i is the variable v_k then $b_i = x_k$, while if t_i is the constant c_ξ, then $b_i = a_\xi$.

Having defined satisfaction for atomic formulas the definition for all formulas of L_β is the same as for L.

Finally we note that for interpretative purposes constants may be replaced by unary predicate letters as is shown by the following construction. We let L^+ be the language obtained from L by adding the unary predicate letters $\{P_\xi : \xi < \beta\}$. Suppose $(\mathfrak{A}, a)$ is a realization of L_β. Then we let

$$\mathfrak{A}_a = \langle A, \{R_n : n \in \omega\} \cup \{X_\xi : \xi < \beta\} \rangle$$

be the realization of L^+ defined by putting $X_\xi = \{a_\xi\}$ for each $\xi < \beta$.

We associate with each formula $\phi = \phi(c_{\xi_1}, \ldots, c_{\xi_n})$ of L_β containing just the constants $c_{\xi_1}, \ldots, c_{\xi_n}$ the formula ϕ^+ of L^+, where

$$\phi^+ = (\forall x_1) \ldots (\forall x_n)[P_{\xi_1}(x_1) \wedge \cdots \wedge P_{\xi_n}(x_n) \to \phi(x_1, \ldots, x_n)],$$

where $x_1, \ldots, x_n$ are any distinct variables not occurring in ϕ. Then it is easy to see that for any $x \in A^\omega$,

$$(\mathfrak{A}, a) \vDash_x \phi \quad \text{iff} \quad \mathfrak{A}_a \vDash_x \phi^+ .$$

EXERCISE 7.1. Verify this last remark and hence deduce that the completeness theorem holds also for the language L_β.

It should be noted that, unlike some authors, we do not insist that the elements of the sequence a should be distinct. Thus we allow the case where in a realization distinct constants have the same value. Therefore it is true that irrefutable sentences of L_β have countable models even when β is an uncountable ordinal.

§ 8. Function symbols

Another way in which we could extend L would be by adding function symbols. Since an n-ary function defined on a set A can be regarded as a special sort of $(n+1)$-ary relation on A, for our purposes we can do without function symbols, and for the sake of economy in future chapters we do without them, even though this sometimes leads to a little unnaturalness in the formulation of various theories. In this section we give a brief account of function symbols so that the reader can restate any subsequent result in a more natural way if he chooses.

We let L′ be the language obtained from L by adding a set $\{f_n : n \in \omega\}$ of

function letters. With each function letter f_n there is associated a positive integer $\gamma(n)$, the *order* of f_n.

The formation rules of L′ are as follows. We first define the notion of a *term* of L′:

(1). A variable standing by itself is a term.

(2). If $t_1,\ldots,t_{\gamma(n)}$ are terms and f_n is a function letter of order $\gamma(n)$, then $f_n(t_1,\ldots,t_{\gamma(n)})$ is a term.

(3). A string is a term only if it follows from a finite number of applications of rules (1) and (2) that it is a term.

Next we define the notion of an *atomic formula* of L′. An atomic formula of L′ is a string which is either of the form

$$t_1 = t_2,$$

or of the form

$$P_n(t_1,\ldots,t_{\delta(n)}),$$

where $t_1,\ldots,t_{\delta(n)}$ are terms.

The notion of a *formula* of L′ is defined recursively in terms of the atomic formulas exactly as for L.

A realization of L′ is a triple

$$\mathfrak{A}' = \langle A, \{R_n : n \in \omega\}, \{F_n : n \in \omega\}\rangle,$$

where $\langle A, \{R_n : n \in \omega\}\rangle$ is a realization of L and for each $n \in \omega$, F_n is a $\gamma(n)$-ary function defined on A.

Before we can define the notion of satisfaction for L′ we need to say what is meant by the *value of the term t in the realization* $\mathfrak{A}'$ *with respect to the valuation* x. We denote this value by $\mathfrak{A}'[t]_x$; it is defined thus:

$$\mathfrak{A}'[v_n]_x = x_n,$$

and

$$\mathfrak{A}'[f_n(t_1,\ldots,t_{\gamma(n)})]_x = F_n(\mathfrak{A}'[t_1]_x,\ldots,\mathfrak{A}'[t_{\gamma(n)}]_x).$$

The notion of *satisfaction* is extended to atomic formulas of L′ in the following way:

$\mathfrak{A}' \vDash_x t_1 = t_2$ iff $\mathfrak{A}'[t_1]_x$, is the same element as $\mathfrak{A}'[t_2]_x$,

$\mathfrak{A}' \vDash_x P_n(t_1,\ldots,t_{\delta(n)})$ iff $\langle \mathfrak{A}'[t_1]_x,\ldots,\mathfrak{A}'[t_{\delta(n)}]_x\rangle \in R_n$.

Then *satisfaction* can be defined for all formulas of L′ just as we did for L in §2.

We say that the term t is *free for* the variable v_n in the formula ϕ if each variable occurring in t is free for v_n in ϕ. The axioms for the extended language are the same as those for L except that in PC8 y can be any term which is free for x in ϕ. The rules of inference are unchanged.

We leave it to the reader to extend the completeness theorem to the language L′ and also to investigate how the definitions of this section should be altered if we add constants to the language L′ with function letters.

§9. Historical and bibliographical remarks.

For the origins of predicate calculus the reader is referred to W. and M. KNEALE [1962] and the historical notes in CHURCH [1956] pages 288–294.

The recursive definition of the notion of satisfaction is due to TARSKI [1935]. The completeness theorem for predicate calculus was first proved by GÖDEL [1930]. We discuss Gödel's proof in ch. 12. Various other proofs have been given since, notably by HENKIN [1949]. The proof using Boolean algebras presented here is due to RASIOWA and SIKORSKI [1951].

Our remarks about the relationship between V and *Val* are based on a remark of Kreisel's (see e.g. KREISEL [1967] page 153, but note that we have reversed Kreisel's notation). The deduction theorem was first stated by HERBRAND [1928].

CHAPTER 4

BEGINNING MODEL THEORY

In this chapter we introduce some of the more fundamental concepts in model theory and we prove some of the basic theorems about models.

§1. The basic notions

In ch. 3 we described the language L of predicate calculus with equality and then we looked at interpretations, i.e. realizations, of this language. We took relational structure as these interpretations. In this chapter and for most of the rest of this book we are going to look at things the other way round. We regard relational structures as the objects of primary interest and we introduce a first order predicate language in order to be able to say things about them. This is the more natural viewpoint in mathematics; it is the structures that present themselves first.

As we see in the next chapter very many mathematical structures can be regarded as relational structures. It may seem that we have lost some generality by only considering sets with relations and not also sets with operations defined on them, but this is not so. Given a set X with an n-ary function f defined on it we define the $(n+1)$-ary relation R_f on X by

$$\langle x_1,\ldots,x_{n+1}\rangle \in R_f \quad \text{iff} \quad f(x_1,\ldots,x_n)=x_{n+1}\,.$$

Clearly the properties of the set with an operator $\langle X,f\rangle$ are precisely determined by the properties of the relational structure $\langle X,\{R_f\}\rangle$. In this way we can regard groups, fields and similar structures as relational structures. Thus the theory of relational structures which we are about to develop has very general applications to large parts of mathematics.

(It should however be noted that there are some areas where, say, the difference between unary operators and binary relations is important. For example, predicate calculus without equality and with only unary predicate letters and unary function letters is decidable, but predicate calculus with

just one binary predicate letter is undecidable. In this book we shall not be concerned with questions where this difference matters.)

Let $\mathfrak{A}=\langle A,\{R_\xi:\xi<\alpha\}\rangle$ be a relational structure and let $\mu\in\omega^\alpha$ be a function defined on α with finite integers as values. We say that $\mathfrak{A}$ is of *type* μ, if, for each $\xi<\alpha$, R_ξ is a $\mu(\xi)$-ary relation on A, that is if $R_\xi\subseteq A^{\mu(\xi)}$.

Given a relational structure $\mathfrak{A}$ of type μ the appropriate language for talking about $\mathfrak{A}$ is predicate calculus with the set $\{P_\xi:\xi<\alpha\}$ of predicate letters, where, for each $\xi<\alpha$, $\delta(\xi)$, the degree of P_ξ, is equal to $\mu(\xi)$. This language will be suitable for talking about any relational structure of type μ. From now on we assume that whenever we use a language to make statements about a relational structure the language is the appropriate one for the particular structure in question.

For the rest of this chapter we assume that all the relational structures we mention are of the same fixed type, say μ. We now consider the various relations that can hold between such structures.

Let

$$\mathfrak{A}=\langle A,\{R_\xi:\xi<\alpha\}\rangle \quad \text{and} \quad \mathfrak{B}=\langle B,\{S_\xi:\xi<\alpha\}\rangle$$

be two relational structures of type μ. We say that $\mathfrak{A}$ is a *substructure* or *subrealization* of $\mathfrak{B}$ and that $\mathfrak{B}$ is an *extension* of $\mathfrak{A}$ if $A\subseteq B$ and each of the relations of $\mathfrak{A}$ is the restriction of the corresponding relation of $\mathfrak{B}$ to A, that is, if for each $\xi<\alpha$

$$R_\xi=S_\xi\cap A^{\mu(\xi)}.$$

If $\mathfrak{A}$ is a substructure of $\mathfrak{B}$ we write $\mathfrak{A}\subseteq\mathfrak{B}$. In particular to each non-empty subset X of B there corresponds the substructure

$$\langle X,\{S_\xi\cap X^{\mu(\xi)}:\xi<\alpha\}\rangle$$

of $\mathfrak{B}$; we call this substructure the *restriction* of $\mathfrak{B}$ to X and denote it by $\mathfrak{B}\restriction X$. In the case of realizations $\mathfrak{A}'=(\mathfrak{A},a)$, $\mathfrak{B}'=(\mathfrak{B},b)$ of a language with constants we say that $\mathfrak{A}'$ is a *substructure* of $\mathfrak{B}'$ if $\mathfrak{A}\subseteq\mathfrak{B}$ and $a=b$.

Let h be a map of A into B. We say that h is a *homomorphism* of $\mathfrak{A}$ into $\mathfrak{B}$ if for each $\xi<\alpha$ and any $a_1,\ldots,a_{\mu(\xi)}$ in A

$$\langle a_1,\ldots,a_{\mu(\xi)}\rangle\in R_\xi \quad \text{iff} \quad \langle h(a_1),\ldots,h(a_{\mu(\xi)})\rangle\in S_\xi.$$

Thus $\mathfrak{A}\subseteq\mathfrak{B}$ iff $A\subseteq B$ and the map $i:A\to B$ defined by $i(a)=a$ is a homomorphism of $\mathfrak{A}$ into $\mathfrak{B}$. i is called the *injection map* of $\mathfrak{A}$ into $\mathfrak{B}$.

A one–one homomorphism of $\mathfrak{A}$ into $\mathfrak{B}$ is called an *isomorphism* between $\mathfrak{A}$ and $\mathfrak{B}\restriction h[A]$. An isomorphism between $\mathfrak{A}$ and $\mathfrak{B}\restriction h[A]$ is also called an *embedding* of $\mathfrak{A}$ into $\mathfrak{B}$. In particular, if h maps A *onto* B, h is an isomorphism

between $\mathfrak{A}$ and $\mathfrak{B}$. If there is an isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$ we say that $\mathfrak{A}$ is isomorphic to $\mathfrak{B}$ and we write $\mathfrak{A} \simeq \mathfrak{B}$. $\simeq$ is an equivalence relation between relational structures. If $\mathfrak{A} \simeq \mathfrak{B}$ then $\mathfrak{A}$ and $\mathfrak{B}$ are essentially the same relational structure.

Isomorphism is a purely algebraic relation involving only the algebraic structure of relational structures and not the language L. The next equivalence relation between relational structures that we consider involves the language L in an essential way. It may happen that although two structures are not isomorphic there is no sentence of the language L that can distinguish between them. In this case we say that they are elementarily equivalent. That is, we say that $\mathfrak{A}$ is *elementarily equivalent* to $\mathfrak{B}$ if each sentence of L true in $\mathfrak{A}$ is also true in $\mathfrak{B}$. If $\mathfrak{A}$ is elementarily equivalent to $\mathfrak{B}$ we write $\mathfrak{A} \equiv \mathfrak{B}$. The following trivial lemma implies that $\equiv$ is an equivalence relation on relational structures.

LEMMA 1.1. *If* $\mathfrak{A} \equiv \mathfrak{B}$ *then for each sentence* σ *of* L

$$\mathfrak{A} \models \sigma \quad \textit{iff} \quad \mathfrak{B} \models \sigma.$$

PROOF. One way round is immediate from the definition. Suppose not $\mathfrak{A} \models \sigma$, then $\mathfrak{A} \models \neg\sigma$. Therefore $\mathfrak{B} \models \neg\sigma$ and so σ is not true in $\mathfrak{B}$.

In a natural sense the properties of relational structures that can be described by sentences of L are simple or elementary. Hence the term *elementarily equivalent* for the notion we have just introduced. Later we shall be interested in finding out which properties of relational structures are elementary in this sense.

EXAMPLES

1.2. If $\mathfrak{A} \simeq \mathfrak{B}$, then $\mathfrak{A} \equiv \mathfrak{B}$.

1.3. Let R be the set of rational numbers and S the set of real numbers, and let $<$ be the usual strict ordering of these sets. Since R is countable while S is uncountable $\langle R, < \rangle$ is not isomorphic to $\langle S, < \rangle$, but as we show later in this section $\langle R, < \rangle \equiv \langle S, < \rangle$.

1.4. Let $\langle A, < \rangle$ be a totally ordered set of order type ω and let $\langle B, < \rangle$ be a totally ordered set of order type $\omega + \omega^* + \omega$. Then $\langle A, < \rangle \equiv \langle B, < \rangle$. This result is proved in ch. 9, §5(f).

If $\mathfrak{A} \subseteq \mathfrak{B}$, then for any *atomic* formula ϕ of L and any valuation $x \in A^\omega$, we have

$$\mathfrak{A} \models_x \phi \quad \text{iff} \quad \mathfrak{B} \models_x \phi.$$

If this equivalence holds not just for atomic formulas but for *any* formula of L we say that $\mathfrak{A}$ is an *elementary substructure* of $\mathfrak{B}$ and that $\mathfrak{B}$ is an *elementary extension* of $\mathfrak{A}$. That is, $\mathfrak{B}$ is an *elementary extension* of $\mathfrak{A}$ if $\mathfrak{A} \subseteq \mathfrak{B}$ and for any formula ϕ of L and any $x \in A^\omega$,

$$\mathfrak{A} \vDash_x \phi \quad \text{iff} \quad \mathfrak{B} \vDash_x \phi .$$

If $\mathfrak{B}$ is an elementary extension of $\mathfrak{A}$ we write $\mathfrak{A} \prec \mathfrak{B}$. Clearly $\mathfrak{A} \prec \mathfrak{B}$ iff $\mathfrak{A} \subseteq \mathfrak{B}$ and for any formula $\phi(v_0,\ldots,v_n)$ of L and any $a_0,\ldots,a_n$ in A,

$$\mathfrak{A} \vDash \phi[a_0,\ldots,a_n] \quad \text{iff} \quad \mathfrak{B} \vDash \phi[a_0,\ldots,a_n] .$$

An embedding h of $\mathfrak{A}$ into $\mathfrak{B}$ is said to be an *elementary embedding* of $\mathfrak{A}$ into $\mathfrak{B}$ if for each formula $\phi(v_0,\ldots,v_n)$ of L and any $a_0,\ldots,a_n$ in A, we have

$$\mathfrak{A} \vDash \phi[a_0,\ldots,a_n] \quad \text{iff} \quad \mathfrak{B} \vDash \phi[h(a_0),\ldots,h(a_n)] .$$

Thus if $\mathfrak{A} \subseteq \mathfrak{B}$, $\mathfrak{A} \prec \mathfrak{B}$ iff the injection of $\mathfrak{A}$ into $\mathfrak{B}$ is an elementary embedding. If there is an elementary embedding of $\mathfrak{A}$ into $\mathfrak{B}$ we say that $\mathfrak{A}$ is *elementarily embeddable* in $\mathfrak{B}$.

The relation $\equiv$ is defined in terms of all sentences while $\prec$ is defined in terms of all formulas. When we want to prove results about all formulas we can use induction on the number of logical symbols in a formula. We have already used this sort of argument. No such argument works for proving results about all sentences. The relation $\prec$ is therefore easier to work with than the relation $\equiv$.

EXERCISE 1.5. Let h be an isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$. Show that h is an elementary embedding of $\mathfrak{A}$ into $\mathfrak{B}$.

EXAMPLE 1.6. Let $\langle R, < \rangle$ and $\langle S, < \rangle$ be as in example 1.3. Then $\langle R, < \rangle \subseteq \langle S, < \rangle$ and we show below that $\langle R, < \rangle \prec \langle S, < \rangle$.

Trivially if $\mathfrak{A} \prec \mathfrak{B}$, then $\mathfrak{A} \equiv \mathfrak{B}$. The next example shows that the converse is not true even when $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{A} \simeq \mathfrak{B}$.

EXAMPLE 1.7. Let Z be the set of non-negative integers and N the set of positive integers and let $<$ be the usual ordering of these sets. Then, if $\mathfrak{A} = \langle N, < \rangle$ and $\mathfrak{B} = \langle Z, < \rangle$, $\mathfrak{A} \subseteq \mathfrak{B}$ and clearly $\mathfrak{A} \simeq \mathfrak{B}$.

Now let $\phi(v_0)$ be the formula

$$\neg(\exists v_1) P(v_1, v_0),$$

then
$$\langle N, < \rangle \models \phi[1],$$
and
$$\langle Z, < \rangle \models \neg\phi[1],$$

when P is interpreted as $<$, since 1 is the least element of N but has a predecessor in Z. Thus, not $\langle N, < \rangle \prec \langle Z, < \rangle$.

The next lemma provides a very convenient criterion for an extension of a relational structure to be an elementary extension.

LEMMA 1.8. *Suppose* $\mathfrak{A} \subseteq \mathfrak{B}$. *Then* $\mathfrak{A} \prec \mathfrak{B}$ *iff whenever* ϕ *is a formula of* L *and* $x \in A^\omega$ *is a sequence of elements of* A *such that* $\mathfrak{B} \models_x (\exists v_n)\phi$, *there is some element* $a \in A$, *such that* $\mathfrak{B} \models_{x(n/a)} \phi$.

PROOF. Suppose first that $\mathfrak{A} \prec \mathfrak{B}$ and $\mathfrak{B} \models_x (\exists v_n)\phi$. Then $\mathfrak{A} \models_x (\exists v_n)\phi$ and so for some $a \in A$, $\mathfrak{A} \models_{x(n/a)} \phi$. $x(n/a) \in A^\omega$ and therefore $\mathfrak{B} \models_{x(n/a)} \phi$.

Conversely suppose that for any formula ϕ of L and any $x \in A^\omega$, $\mathfrak{B} \models_x (\exists v_n)\phi$ implies that there is some $a \in A$ such that $\mathfrak{B} \models_{x(n/a)} \phi$. Then we show that $\mathfrak{A} \prec \mathfrak{B}$, that is, we prove that

$$\text{for any formula } \phi \text{ of L and any } x \in A^\omega,\ \mathfrak{A} \models_x \phi \quad \text{iff} \quad \mathfrak{B} \models_x \phi. \qquad [*]$$

We prove [*] by induction on the number of logical symbols in ϕ. It is trivially true when ϕ is an atomic formula since $\mathfrak{A} \subseteq \mathfrak{B}$. Clearly if [*] holds for the formulas ψ and χ then it holds also for $\neg\psi$ and for $\psi \wedge \chi$.

Suppose [*] holds for the formula ψ. We will prove that it holds also for $(\exists v_n)\psi$. If $\mathfrak{A} \models_x (\exists v_n)\psi$ there is some $a \in A$ such that $\mathfrak{A} \models_{x(n/a)} \psi$. By hypothesis [*] holds for ψ. Therefore $\mathfrak{B} \models_{x(n/a)} \psi$ and so $\mathfrak{B} \models_x (\exists v_n)\psi$. If $\mathfrak{B} \models_x (\exists v_n)\psi$, by the hypothesis of the lemma there is some $a \in A$ such that $\mathfrak{B} \models_{x(n/a)} \psi$. By the induction hypothesis, $\mathfrak{A} \models_{x(n/a)} \psi$ and therefore $\mathfrak{A} \models_x (\exists v_n)\psi$. This completes the proof.

COROLLARY 1.9. *If* $\mathfrak{A} \subseteq \mathfrak{B}$, *then* $\mathfrak{A} \prec \mathfrak{B}$ *iff whenever* $\phi(v_0, \ldots, v_n)$ *is a formula of* L *with free variables among* $v_0, \ldots, v_n$, *and* $a_0, \ldots, a_{n-1}$ *are elements of* A *such that for some* $b \in B$, $\mathfrak{B} \models \phi[a_0, \ldots, a_{n-1}, b]$, *then there is some* $a \in A$ *such that* $\mathfrak{B} \models \phi[a_0, \ldots, a_{n-1}, a]$.

PROOF. This is just a restatement of the lemma.

We immediately use this corollary to show that the reals considered as an ordered set are an elementary extension of the rationals, that is, in the notation of example 1.3, that $\langle R, < \rangle \prec \langle S, < \rangle$. So suppose ϕ is a formula of L (since we are dealing in this case with relational systems with a single binary relation, L has just one predicate letter of degree 2) with free vari-

ables among $v_0,\ldots,v_n$ and that $a_0,\ldots,a_{n-1}$ are rationals and b is a real number such that

$$\langle S,<\rangle \models \phi[a_0,\ldots,a_{n-1},b].$$

Without loss of generality we can assume

$$a_0<a_1<\cdots<a_{n-1}.$$

If b is a rational there is nothing to prove. Suppose then that b is irrational, say

$$a_k<b<a_{k+1} \quad \text{with} \quad k+1\leqslant n-1.$$

(The trivial modification of the argument in case $b<a_0$ or $a_{n-1}<b$ is left to the reader.)

Let c be any rational between a_k and a_{k+1}. There is an order isomorphism h of S onto S which leaves each a_i fixed and maps b onto c. To be specific, we let

$$h(x)=x, \qquad \text{if } x\leqslant a_k \text{ or } a_{k+1}\leqslant x,$$

$$h(x)=\frac{c-a_k}{b-a_k}(x-a_k)+a_k, \qquad \text{if } a_k\leqslant x\leqslant b,$$

and

$$h(x)=\frac{a_{k+1}-c}{a_{k+1}-b}(x-b)+c, \qquad \text{if } b\leqslant x\leqslant a_{k+1},$$

then h has the stated properties. In particular h is an elementary embedding of $\langle S,<\rangle$ onto $\langle S,<\rangle$ and so

$$\langle S,<\rangle \models \phi[h(a_0),\ldots,h(a_{n-1}),h(b)],$$

that is

$$\langle S,<\rangle \models \phi[a_0,\ldots,a_{n-1},c].$$

But c is a rational and so is in R. Therefore by corollary 1.9., we have shown that $\langle R,<\rangle\prec\langle S,<\rangle$. We can also conclude that $\langle R,<\rangle\equiv\langle S,<\rangle$. Another proof of this result will be given in ch. 9.

Other useful criteria for an extension to be an elementary extension are given by the next theorem.

THEOREM 1.10. *If* $\mathfrak{A}\subseteq\mathfrak{B}$, *the following conditions are equivalent:*

(1). $\mathfrak{A}\prec\mathfrak{B}$,

(2). *for each* $a\in A^\beta$, *with* $\beta<\omega$, $(\mathfrak{A},a)\equiv(\mathfrak{B},a)$,

(3). *for each* $a\in A^\beta$, *with* β *arbitrary*, $(\mathfrak{A},a)\prec(\mathfrak{B},a)$,

(4). *for each* $a\in A^\beta$, *with* β *arbitrary*, $(\mathfrak{A},a)\equiv(\mathfrak{B},a)$.

(In conditions (2), (3) and (4) the relations $\equiv$, $\prec$ are to be understood of course as referring to the language L_β obtained from L by adding the β-termed sequence of constants $\langle c_\xi : \xi < \beta \rangle$.)

PROOF. Assume (1). Suppose β is an arbitrary ordinal and that $a \in A^\beta$. Suppose ϕ is a formula of L_β and $b \in A^\omega$ is such that $(\mathfrak{A},a) \models_b \phi$. For simplicity we assume ϕ has just the one constant c_ξ. Let v_k be a variable which does not occur in ϕ, and let ϕ' be the formula which results from ϕ by replacing all occurrences of c_ξ by v_k. Clearly $(\mathfrak{A},a) \models_{b(k/a_\xi)} \phi'$. ϕ' contains no constants and is therefore a sentence of L. Therefore $\mathfrak{A} \models_{b(k/a_\xi)} \phi'$. Since $\mathfrak{A} \prec \mathfrak{B}$ it follows that $\mathfrak{B} \models_{b(k/a_\xi)} \phi'$ and hence that $(\mathfrak{B},a) \models_b \phi$. We have thus shown that (3) holds.

Trivially (3) implies (4) and (4) implies (2), so it remains only to show that (2) implies (1).

Assume (2). Let ϕ be a formula of L and let $b \in A^\omega$ be such that $\mathfrak{A} \models_b \phi$. Replace all the distinct variables $v_{i_1},\ldots,v_{i_n}$ which occur free in ϕ by the constants $c_1,\ldots,c_n$, respectively. The resulting formula ϕ' is a sentence of L_{n+1}. Let b' be the sequence $\langle b_{i_1},\ldots,b_{i_n} \rangle$. Then clearly $(\mathfrak{A},b') \models \phi'$ and so by (2), $(\mathfrak{B},b') \models \phi'$ and from this it follows that $\mathfrak{B} \models_b \phi$. Thus we have proved (1).

Let A be any set. A sequence $a \in A^\alpha$, where α is some ordinal, is called an *enumeration* of A if each element of A occurs at least once in the sequence a. a is called an *enumeration of A without repetitions* if each element of A occurs precisely once in the sequence a.

Subsequently we make use of the following results.

THEOREM 1.11. *Let $\mathfrak{A}$ be a realization of* L *and let $a \in A^\alpha$ be an enumeration of A. If $\mathfrak{B}$ is also a realization of* L, *$\mathfrak{A}$ is elementarily embeddable in $\mathfrak{B}$ iff there is a sequence $b \in B^\alpha$ such that, with respect to the extended language* L_α,

$$(\mathfrak{A},a) \equiv (\mathfrak{B},b).$$

PROOF. Suppose first that h is an elementary embedding of $\mathfrak{A}$ into $\mathfrak{B}$. If we define the sequence $b \in B^\alpha$ by putting, for each $\xi < \alpha$,

$$b_\xi = h(a_\xi),$$

then it is clear that $(\mathfrak{A},a) \equiv (\mathfrak{B},b)$.

Conversely, suppose that there is a sequence $b \in B^\alpha$ such that $(\mathfrak{A},a) \equiv (\mathfrak{B},b)$. If we define the map $h : A \to B$ by putting, for each $\xi < \alpha$,

$$h(a_\xi) = b_\xi,$$

then it is equally clear that h is an elementary embedding of $\mathfrak{A}$ into $\mathfrak{B}$.

THEOREM 1.12. *Let* $\mathfrak{A}$, $\mathfrak{B}$ *be two realizations of* L *of the same cardinality and let* $a \in A^{\alpha}$ *be an enumeration of* A. *Then*

(1) *if* $\mathfrak{A} \simeq \mathfrak{B}$, *there is some* $b \in B^{\alpha}$ *such that* $(\mathfrak{A},a) \equiv (\mathfrak{B},b)$;

(2) *if* b *is an enumeration of* B *and* $(\mathfrak{A},a) \equiv (\mathfrak{B},b)$, *then* $\mathfrak{A} \simeq \mathfrak{B}$;

(3) *if* $A \subseteq B$ *then* $\mathfrak{A} \prec \mathfrak{B}$ *iff* $(\mathfrak{A},a) \equiv (\mathfrak{B},a)$.

PROOF. This is evident from the previous theorem and its proof.

EXERCISE 1.13. Write out a detailed proof of theorem 1.12.

§ 2. Unions of chains

A sequence $\langle \mathfrak{A}_{\xi} : \xi < \alpha \rangle$ is said to be a *chain* of relational structures, if whenever $\xi \leqslant \eta < \alpha$ we have $\mathfrak{A}_{\xi} \subseteq \mathfrak{A}_{\eta}$. Suppose that for $\xi < \alpha$,

$$\mathfrak{A}_{\xi} = \langle A_{\xi}, \{R_n^{\xi} : n \in \omega\} \rangle$$

By the *union* of the chain $\langle \mathfrak{A}_{\xi} : \xi < \alpha \rangle$ which we denote by $\bigcup_{\xi<\alpha} \mathfrak{A}_{\xi}$, we mean the relational system

$$\langle \bigcup_{\xi<\alpha} A_{\xi}, \{ \bigcup_{\xi<\alpha} R_n^{\xi} : n \in \omega \} \rangle .$$

$\bigcup_{\xi<\alpha} \mathfrak{A}_{\xi}$ is an extension of each of the structures $\mathfrak{A}_{\xi}$ but in general its elementary properties can be different from those of the elements of the chain. For example if for each $n \in \omega$, $\mathfrak{A}_n = \langle A_n, < \rangle$ where $A_n = \{0,1,\ldots,n\}$ and $<$ is the usual order relation on this set, then $\langle \mathfrak{A}_n : n \in \omega \rangle$ is a chain of relational systems. The sentence which says that the ordering $<$ has a greatest element, namely the sentence $(\exists v_0)(\forall v_1)[v_0 = v_1 \vee P(v_1,v_0)]$, is true in each $\mathfrak{A}_n$, but is not true in the union $\bigcup_{n \in \omega} \mathfrak{A}_n$.

We shall be more interested in the following case when the elementary properties of the union are determined by those of the elements of the chain. The chain $\langle \mathfrak{A}_{\xi} : \xi < \alpha \rangle$ is called an *elementary chain* if $\mathfrak{A}_{\zeta} \prec \mathfrak{A}_{\xi}$ whenever $\zeta \leqslant \xi < \alpha$

THEOREM 2.1. UNION OF CHAINS

The union of an elementary chain is an elementary extension of each element of the chain.

PROOF. Let $\langle \mathfrak{A}_{\xi} : \xi < \alpha \rangle$ be an elementary chain whose union is $\mathfrak{A}$. We prove that

for any formula $\phi(v_0,\ldots,v_n)$, any $\xi < \alpha$ and any $a_0,\ldots,a_n$ in A_{ξ},

$$\mathfrak{A} \models \phi[a_0,\ldots,a_n] \text{ iff } \mathfrak{A}_{\xi} \models \phi[a_0,\ldots,a_n]. \qquad [*]$$

The proof of [*] is by induction on the number of logical symbols in ϕ. Since $\mathfrak{A}$ is an extension of each $\mathfrak{A}_{\xi}$, it is trivial for atomic formulas. Clearly

if it holds for the formulas ϕ and ψ then it holds also for $\neg\phi$ and $\phi \wedge \psi$.

Suppose $[*]$ holds for the formula $\psi(v_0,\ldots,v_n)$. We show that it holds also for $\phi(v_0,\ldots,v_{n-1}) = (\exists v_n)\,\psi(v_0,\ldots,v_n)$. If $a_0,\ldots,a_{n-1} \in A_\xi$ and $\mathfrak{A}_\xi \models \phi[a_0,\ldots,a_{n-1}]$ then for some $a_n \in A_\xi$, $\mathfrak{A}_\xi \models \psi[a_0,\ldots,a_n]$. Therefore, by the induction hypothesis, $\mathfrak{A} \models \psi[a_0,\ldots,a_n]$, whence $\mathfrak{A} \models \phi[a_0,\ldots,a_{n-1}]$.

Conversely, assume that $\mathfrak{A} \models \phi[a_0,\ldots,a_{n-1}]$. Then $\mathfrak{A} \models \psi[a_0,\ldots,a_{n-1},a]$, for some $a \in A$. For some $\eta < \alpha$, $a \in A_\eta$. Let ζ be the greater of ξ and η. Then $a_0,\ldots,a_{n-1},a \in A_\zeta$ and so, by the induction hypothesis, $\mathfrak{A}_\zeta \models \psi[a_0,\ldots,a_{n-1},a]$, whence $\mathfrak{A}_\zeta \models \phi[a_0,\ldots,a_{n-1}]$. But $\xi \leqslant \zeta$ and so because $\mathfrak{A}_\xi \prec \mathfrak{A}_\zeta$ we can conclude that $\mathfrak{A}_\xi \models \phi[a_0,\ldots,a_{n-1}]$, completing the proof.

The usefulness of this result will become apparent later.

§ 3. Löwenheim–Skolem theorems

In this section we shall be concerned with the following question. Suppose $\mathfrak{A}$ is a relational structure of cardinal α, then for which cardinals β does $\mathfrak{A}$ have elementary extensions and substructures of cardinal β?

By the *cardinal* of a language L we mean the cardinality of the set of symbols of L. Since L has an infinite number of variables the cardinal of L is always infinite. Formulas of L are *finite* strings of symbols of L. Therefore the cardinality of the set of formulas of L is the same as the cardinal of L.

Throughout this section we assume that all relational structures are of the same type μ, and that the language L appropriate for structures of this type has cardinal ρ.

THEOREM 3.1. *Let $\mathfrak{A}$ be an infinite relational structure of cardinal α, and let C be a subset of A of cardinal γ. If β is a cardinal satisfying $\gamma, \rho \leqslant \beta \leqslant \alpha$ then $\mathfrak{A} \restriction C$ has an extension $\mathfrak{B}$ of cardinal β which is an elementary substructure of $\mathfrak{A}$.*

PROOF. Let $<$ be some fixed well-ordering of A. By 'least' we will mean 'least in the ordering $<$'. We define a sequence $\langle B_n : n < \omega \rangle$ of subsets of A as follows:

B_0 is any subset of A of cardinal β which contains C, and in general B_{n+1} is the set of all elements $a \in A$ such that for some formula $\phi(v_0,\ldots,v_n)$ of L, and some $a_0,\ldots,a_{n-1}$ in B_n, a is the least element of A such that

$$\mathfrak{A} \models \phi[a_0,\ldots,a_{n-1},a].$$

For each $a_0 \in B_n$, a_0 is the least (indeed the only) element a of A such that $\mathfrak{A} \models v_0 = v_1[a_0,a]$. Therefore $a_0 \in B_{n+1}$ and so $B_n \subseteq B_{n+1}$.

Since there are only ρ distinct formulas of L and $\rho \leqslant \beta$ each set B_n is of cardinal β. Therefore $B = \bigcup_{n\in\omega} B_n$ is of cardinal β. Let $\mathfrak{B} = \mathfrak{A} \restriction B$. Since $C \subseteq B$, $\mathfrak{B}$ is an extension of $\mathfrak{A} \restriction C$. We show that $\mathfrak{B} \prec \mathfrak{A}$.

Suppose $\phi(v_0,\ldots,v_n)$ is a formula of L, $a_0,\ldots,a_{n-1}$ are in $\mathfrak{B}$ and there is some $a \in A$ such that $\mathfrak{A} \models \phi[a_0,\ldots,a_{n-1},a]$.

For $0 \leqslant i < n$, $a_i \in B$ and so there is some $n_i < \omega$ with $a_i \in B_{n_i}$. Let m be the maximum of the set $\{n_i : 0 \leqslant i < n\}$. Then $a_0,\ldots,a_{n-1} \in B_m$. By hypothesis the set $\{a \in A : \mathfrak{A} \models \phi[a_0,\ldots,a_{n-1},a]\}$ is not empty and so contains a least element say b. By construction $b \in B_{m+1} \subseteq B$. Thus there is some $b \in B$, such that $\mathfrak{A} \models \phi[a_0,\ldots,a_{n-1},b]$. Since $\mathfrak{B} \subseteq \mathfrak{A}$ it follows from corollary 1.9 that $\mathfrak{B} \prec \mathfrak{A}$.

COROLLARY 3.2. THE DOWNWARD LÖWENHEIM–SKOLEM THEOREM
Let Σ be a set of sentences of cardinal α, with an infinite model of cardinal $\beta \geqslant \alpha$. Then Σ has a model of any infinite *cardinal γ such that $\alpha \leqslant \gamma \leqslant \beta$.*

PROOF. Let $\delta = \max\{\aleph_0, \alpha\}$. At most δ predicate letters occur in Σ and so we may regard Σ as a set of sentences of a language L′ of cardinal δ.

Let $\mathfrak{A}$ be the model of Σ of cardinal $\beta \geqslant \alpha$ which, by hypothesis, exists. By dropping the relations corresponding to the predicate letters not occurring in L′, we obtain a realization $\mathfrak{A}'$ of cardinal β which is a model of Σ. If γ is an infinite cardinal such that $\alpha \leqslant \gamma \leqslant \beta$, then $\delta \leqslant \gamma \leqslant \beta$, and so by the previous theorem $\mathfrak{A}'$ has an elementary substructure $\mathfrak{B}'$ of cardinal γ. Since $\mathfrak{B}' \prec \mathfrak{A}'$, $\mathfrak{B}'$ is also a model of Σ. By adding arbitrary relations to $\mathfrak{B}'$ as values of the predicate letters not occurring in Σ, we obtain a realization $\mathfrak{B}$ of L of cardinal γ which is a model of Σ.

Next we show that a given infinite relational structure has elementary extensions of arbitrarily large cardinality. To prove this we shall need to assume the truth of the following theorem whose proof is deferred until the next chapter (see theorem 5.4.1).

THE COMPACTNESS THEOREM FOR PREDICATE CALCULUS
A set Σ of sentences of L *has a model iff each finite subset of Σ has a model.*

THEOREM 3.3. *Let $\mathfrak{A}$ be an infinite realization of* L *of cardinal α. $\mathfrak{A}$ has an elementary extension of any cardinal $\beta \geqslant \alpha, \rho$.*

PROOF. Suppose $\beta \geqslant \alpha$ and $\beta \geqslant \rho$. Let $a \in A^{\alpha}$ be an enumeration of A. Let L_{α} be the language obtained from L by adding the sequence of constants $\langle c_{\xi} : \xi < a \rangle$ and let $L_{\alpha+\beta}$ be the language obtained from L_{α} by adding the constants $\langle d_{\xi} : \xi < \beta \rangle$.

Let Δ be the set of all sentences of L_α which hold in $(\mathfrak{A},a)$ and let $\Gamma=\{\neg d_\xi=d_\zeta:\xi<\zeta<\beta\}$. We will show that each finite subset of $\Delta\cup\Gamma$ has a model. To do this it will be sufficient to show that if Γ_0 is a finite subset of Γ then $\Delta\cup\Gamma_0$ has a model.

Suppose that Γ_0 is a finite subset of Γ. Let $d_{\xi_1},\ldots,d_{\xi_n}$ be the constants occurring in Γ_0. Let $a'_1,\ldots,a'_n$ be any n distinct elements of A (these exist since A is infinite). We define the sequence $a''\in A^{\alpha+\beta}$ as follows:

$a''_\xi=a_\xi$, for $\xi<\alpha$,

$a''_\xi=a'_i$, if $\xi=\alpha+\xi_i$, with $1\leqslant i\leqslant n$,

and otherwise a''_ξ is arbitrary.

Since the a'_i are distinct $(\mathfrak{A},a'')$ is a model of $\Delta\cup\Gamma_0$. Thus we have shown that each finite subset of $\Delta\cup\Gamma$ has a model. We can therefore conclude from the compactness theorem that $\Delta\cup\Gamma$ itself has a model. Since the constants d_ξ have distinct values in any model of $\Delta\cup\Gamma$ this model must be of cardinal $\geqslant\beta$. Therefore, by corollary 3.2, $\Delta\cup\Gamma$ has a model, say $(\mathfrak{B},b)$, of cardinal β. Since Δ is the set of *all* sentences of L_α true in $(\mathfrak{A},a)$, it follows that

$$(\mathfrak{A},a)\equiv(\mathfrak{B},b'),$$

where $b'=b\restriction\alpha$, i.e. b' is the sequence of the first α terms of the sequence b. Since a is an enumeration of A, $\mathfrak{A}$ is elementarily embeddable in $\mathfrak{B}$. Thus, up to isomorphism, $\mathfrak{B}$ is an elementary extension of $\mathfrak{A}$ of cardinal β.

COROLLARY 3.4. THE UPWARD LÖWENHEIM–SKOLEM THEOREM
Let Σ be a set of sentences of cardinal α. If Σ has a model of cardinal $\beta\geqslant\aleph_0$, then Σ has a model of any cardinal $\geqslant\alpha,\beta$.

Putting these results together we obtain

THEOREM 3.5. THE LÖWENHEIM–SKOLEM THEOREM
Let Σ be a set of sentences with an infinite model. If Σ is finite Σ has a model of each infinite cardinal. If Σ is infinite and of cardinal α then Σ has a model of any cardinal $\geqslant\alpha$.

Notice that we used the axiom of choice in our proof of theorem 3.1. right at the beginning in assuming that A could be well-ordered. It is natural to ask whether this use of the axiom of choice is essential. We will show that it is, by proving that the downward Löwenheim–Skolem theorem implies the axiom of choice.

We need first the fact that *without assuming the axiom of choice* we can

show that there are arbitrarily large cardinals α such that $\alpha^2=\alpha$. This is given by the next lemma.

LEMMA 3.6. *For each cardinal α there exists a cardinal β such that $\alpha<\beta$ and $\beta^2=\beta$. This does not need the axiom of choice.*

PROOF. Since we need to be sure that the axiom of choice is not used we spell out the proof in more detail than would otherwise be necessary.

We leave it to the reader to verify that the following facts can all be established without using the axiom of choice:

For all cardinals γ,δ,

(a). $\gamma\leqslant\gamma\cdot\aleph_0$,

(b). $2^\gamma\leqslant 2^\delta$ if $\gamma\leqslant\delta$,

(c). $\gamma<2^\gamma$,

(d). $(\gamma^\delta)^2=\gamma^{\delta\cdot 2}$,

(e). $\gamma\cdot\aleph_0\cdot 2=\gamma\cdot\aleph_0$.

Now suppose α is any cardinal. Let $\beta=2^{\alpha\cdot\aleph_0}$. From (a), (b) and (c) it follows that

$$\alpha<2^\alpha\leqslant 2^{\alpha\cdot\aleph_0}=\beta,$$

and from (d) and (e) that

$$\beta^2=(2^{\alpha\cdot\aleph_0})^2=2^{\alpha\cdot\aleph_0\cdot 2}=2^{\alpha\cdot\aleph_0}=\beta.$$

This concludes the proof.

THEOREM 3.7. *The downward Löwenheim–Skolem theorem implies the axiom of choice.*

PROOF. Assume that the downward Löwenheim–Skolem theorem holds. Let α be an infinite cardinal; we will prove, without using the axiom of choice, that $\alpha^2=\alpha$.

By the previous lemma there is a cardinal β such that without using the axiom of choice we can prove that $\alpha<\beta$ and that $\beta^2=\beta$.

Let σ be the sentence

$$(\forall x)(\forall y)(\exists z)(\forall t)[P(x,y,t)\leftrightarrow z=t]$$
$$\wedge(\forall z)(\exists u)(\exists v)(\forall x)(\forall y)[P(x,y,z)\leftrightarrow x=u\wedge y=v].$$

'P is a predicate which indexes the collection of all ordered pairs of elements of the domain by elements of the domain.'

Since $\beta^2=\beta$, there is a one–one map ϕ from $\beta\times\beta$ onto β. Clearly if $R=\{\langle x,y,z\rangle\in\beta^3:\phi(x,y)=z\}$, then $\mathfrak{A}=\langle\beta,R\rangle$ is a model of σ. By the downward Löwenheim–Skolem theorem it follows that σ has a model of cardinal

α, say $\mathfrak{B}=\langle B,S\rangle$. We define $\psi: B\times B\to B$ by

$$\psi(x,y)=z \quad \text{iff} \quad \langle x,y,z\rangle\in S.$$

Since $\mathfrak{B}$ is a model of σ, ψ is one–one onto. Therefore $\text{card}(B\times B)=\text{card}(B)$, that is, $\alpha^2=\alpha$.

Thus we have shown that, *for all infinite cardinals* α, $\alpha^2=\alpha$. But it is well known (see e.g. SIERPINSKI [1965] p. 422 or RUBIN and RUBIN [1963] p. 52) that this last statement is equivalent to the axiom of choice. Thus our proof of theorem 3.7 is completed.

Essentially our argument was as follows: without the axiom of choice we can prove that there are arbitrarily large cardinals β such that $\beta^2=\beta$. The downward Löwenheim-Skolem theorem enables us to deduce that if $\beta^2=\beta$ then for all infinite cardinals $\alpha\leqslant\beta$, $\alpha^2=\alpha$. Thus $\alpha^2=\alpha$ for all infinite cardinals and this is equivalent to the axiom of choice. A similar argument shows that the upward Löwenheim–Skolem theorem also implies the axiom of choice.

THEOREM 3.8. *The upward Löwenheim–Skolem theorem implies the axiom of choice.*

PROOF. Without the axiom of choice we can prove that $\aleph_0\times\aleph_0=\aleph_0$. Indeed the function f defined by

$$f(m,n)=\tfrac{1}{2}[(m+n)^2+3m+n],$$

is a one–one map of $\omega\times\omega$ onto ω. Hence the sentence σ of theorem 3.7 has a model of cardinal $\aleph_0$. Therefore, by the upward Löwenheim–Skolem theorem, σ has a model of each infinite cardinal α. Therefore, as before, we can draw the conclusion that for each infinite cardinal α, $\alpha^2=\alpha$ and hence that the axiom of choice holds.

§ 4. Hanf numbers

In this section we make a short digression about general formal languages and we consider the possibility of generalizing the Löwenheim–Skolem theorems of the previous section to these languages. We shall see that suitable generalizations hold for any (not necessarily first order) language which satisfies certain weak conditions. Throughout this section we distinguish between *sets* and *proper classes*, using these terms in the sense of GÖDEL [1940].

Let L* be some *formal language*; L* will have certain *basic symbols* and

will have *formation rules* telling us which strings of these symbols count as *sentences*. We suppose also that we have defined the notion of a structure being a *realization* of L* and that of a realization being a *model* of a sentence or set of sentences of L*.

Suppose that there is some cardinal α such that any set of sentences of L* which has a model of cardinal at least α has models of arbitrarily high cardinality. We call the least cardinal with this property the *Hanf number* of L* and denote it by χ_{L^*}. HANF [1960] proved the following existence theorem about such numbers.

THEOREM 4.1. HANF'S THEOREM
If the class of all sentences of L* *forms a set then* L* *has a Hanf number.*

PROOF. Let $\mathscr{S}$ be the collection of all those sets Σ of sentences of L* which do *not* have models of arbitrarily high cardinality. For each $\Sigma \in \mathscr{S}$ there is some cardinal κ_Σ such that all models of Σ are of cardinal less than κ_Σ.

Since the collection of all sentences of L* is a set, $\mathscr{S}$ is also a set and hence there is a cardinal β which is the supremum of the *set* of cardinals $\{\kappa_\Sigma : \Sigma \in \mathscr{S}\}$. Any set Σ of sentences of L* having a model of cardinal at least β has models of arbitrarily high cardinality, since such a set Σ cannot be in $\mathscr{S}$. It follows that L* has a Hanf number.

This result generalizes the upward Löwenheim–Skolem theorem. In a similar fashion one can prove

THEOREM 4.2. *Let Σ be a set of sentences of* L*. *There is a cardinal α such that if a sentence $\sigma \in \Sigma$ has a model of cardinal at least α, then σ has models of arbitrarily high cardinality. The least such α is called the Hanf number of Σ and is denoted by $\chi_{L^*}(\Sigma)$.*

By the *Löwenheim number* of a language L* we shall mean the least cardinal α such that each set of sentences of L* with a model has a model of cardinal $\leqslant \alpha$, if such a cardinal exists. In fact an argument similar to that used above shows that all reasonable languages have a Löwenheim number.

THEOREM 4.3. *If the collection of sentences of* L* *forms a set then* L* *has a Löwenheim number.*

PROOF. Let $\mathscr{S}$ be the collection of all sets of sentences of L* which have a model. If $\Sigma \in \mathscr{S}$ let κ_Σ be the least cardinal such that Σ has a model of cardinal κ_Σ. Since the collection of all sentences of L* is a set, $\mathscr{S}$ is a set. Hence there is a cardinal α which is the supremum of the set $\{\kappa_\Sigma : \Sigma \in \mathscr{S}\}$.

It is easily seen that any set of sentences of L* with a model has a model of cardinal $\leqslant \alpha$ and that α is the Löwenheim number of L*.

This result generalizes the downward Löwenheim–Skolem theorem. In spite of the fact that most languages can easily be shown to possess Hanf numbers, the problem of computing (or even estimating) the Hanf number of a given language is in general very difficult. However for first order predicate languages the calculation is simple. Indeed it follows from the results of the previous section that if L has α predicate letters with α infinite, then the Hanf number of L is $\aleph_0$ and its Löwenheim number is α.

§ 5. Historical and bibliographical remarks

Most of the basic model theoretic ideas introduced in this chapter are due to Tarski, as also are most of the theorems (see especially TARSKI and VAUGHT [1957]). The Löwenheim–Skolem theorem has a complicated history; our account is based on that in VAUGHT [1954]. LÖWENHEIM [1915] proved that a finitely axiomatizable theory with a model has a countable model. This result was extended to arbitrary countable theories by SKOLEM [1920]. The general result for countable theories is due to Tarski (see the editors' remarks at the end of SKOLEM [1934]), and its extension to uncountable theories is due to HENKIN [1949], though in a weaker form it follows from an earlier result of MALCEV [1936].

The remark that both the upward and downward Löwenheim–Skolem theorems are equivalent to the axiom of choice is due to VAUGHT [1956]. Hanf numbers were introduced in HANF [1960].

CHAPTER 5

INTRODUCING ULTRAPRODUCTS

In this chapter we describe the construction that will occupy our attention for most of the rest of the book, namely the ultraproduct construction.

§1. The ultraproduct construction

For ease of exposition we confine our attention to relational structures consisting of a non-empty set and a single binary relation on that set. The appropriate language L for such structures thus has a single predicate letter of degree 2, say P_0. The type of these structures is $\mu_0 = \{\langle 0,2\rangle\}$. It will be clear that everything we do can be extended to relational structures with arbitrarily many finitary relations merely by complicating the notation. We leave it to the reader to make these extensions for himself.

Let I be some arbitrary fixed index set. By an *ultrafilter on* I we mean an ultrafilter in the power set Boolean algebra $\langle \mathscr{P}(I), \subseteq \rangle$.

For each $i \in I$ let $\mathfrak{A}_i = \langle A_i, R_i \rangle$ be a relational structure of type μ_0. Let $A = \prod_{i \in I} A_i$ be the cartesian product of the sets A_i. (Where the index set is clear from the context, we will in future write just $\prod_i A_i$ or $\prod A_i$, and similarly for other expressions of the same sort.)

We will use f,g,h,f',g',h' etc. to denote elements of A. If $f \in A$ we denote its i-th coordinate by either $f(i)$ or f_i.

Let F be a collection of subsets of I. We define the relation $\sim_F$ on A by

$$f \sim_F g \quad \text{iff} \quad \{i \in I : f(i) = g(i)\} \in F .$$

The following result is fundamental.

LEMMA 1.1. *If F is a filter on I, then $\sim_F$ is an equivalence relation on $\prod A_i$.*

PROOF. If F is a filter on I, $I \in F$ and so $\sim_F$ is reflexive. Clearly $\sim_F$ is symmetric; we show that it is also a transitive relation.

Suppose $f \sim_F g$ and $g \sim_F h$, then by the definition of $\sim_F$,

$$X = \{i \in I : f(i) = g(i)\} \in F,$$

and

$$Y = \{i \in I : g(i) = h(i)\} \in F.$$

Therefore, because F is a filter, $X \cap Y \in F$. But

$$X \cap Y \subseteq Z = \{i \in I : f(i) = h(i)\},$$

and so $Z \in F$, that is, $f \sim_F h$. Thus $\sim_F$ is transitive.

EXERCISE 1.2. Prove that the converse of lemma 1.1 is also true, provided that each A_i contains more than two elements and construct a counter-example in the case where some of the A_i contain only two elements.

From now on we will assume that F is a filter on I. Intuitively a filter on I consists of the 'big' subsets of I and we can think of $f \sim_F g$ as saying that f agrees with g at 'almost all' its coordinates. We can extend this idea and define the relation S on $\prod A_i$ by stipulating that S is to hold in $\prod A_i$ whenever R_i holds at 'almost all' the coordinates. The precise definition of S is:

$$\langle f, g \rangle \in S \quad \text{iff} \quad \{i \in I : \langle f(i), g(i) \rangle \in R_i\} \in F.$$

It can be seen that the definition of $\sim_F$ is just the special case where R_i is the identity relation on A_i. The next lemma demonstrates another very important property of filters. It shows that elements of $\prod A_i$ which are equivalent under the relation $\sim_F$ are also indistinguishable as far as the relation S is concerned.

LEMMA 1.3. *$\sim_F$ is a congruence relation with respect to S, that is, if $f \sim_F f'$ and $g \sim_F g'$, then $\langle f, g \rangle \in S$ implies $\langle f', g' \rangle \in S$.*

PROOF. Suppose $f \sim_F f'$ and $g \sim_F g'$. Let

$$X = \{i \in I : f(i) = f'(i)\} \quad \text{and} \quad Y = \{i \in I : g(i) = g'(i)\},$$

then $X, Y \in F$. If $\langle f, g \rangle \in S$,

$$Z = \{i \in I : \langle f(i), g(i) \rangle \in R_i\} \in F.$$

Since F is a filter $X \cap Y \cap Z \in F$. But

$$X \cap Y \cap Z \subseteq W = \{i \in I : \langle f'(i), g'(i) \rangle \in R_i\},$$

and therefore $W \in F$. We can conclude from this that $\langle f', g' \rangle \in S$.

For each $f \in \prod A_i$ we let f/F be the equivalence class to which f belongs under the relation $\sim_F$, and we let

$$\prod A_i/F = \{f/F : f \in \prod A_i\}.$$

The relation S on $\prod A_i$ induces the relation R_F on $\prod A_i/F$ which is defined by

$$\langle f/F, g/G\rangle \in R_F \quad \text{iff} \quad \langle f, g\rangle \in S.$$

By the previous lemma this is a sound definition.

We let $\prod \mathfrak{A}_i/F$ be the relational structure $\langle \prod A_i/F, R_F\rangle$. $\prod \mathfrak{A}_i/F$ is called the *reduced product* of the family $\{\mathfrak{A}_i : i \in I\}$ over the filter F. If F is an ultrafilter $\prod \mathfrak{A}_i/F$ is called an *ultraproduct*. If for each $i \in I$, $\mathfrak{A}_i = \mathfrak{A}$, the reduced product is denoted by $\mathfrak{A}^I/F$ and is called a *reduced power* of $\mathfrak{A}$ over F. If F is an ultrafilter $\mathfrak{A}^I/F$ is called an *ultrapower* of $\mathfrak{A}$.

EXERCISE 1.4. If F is the filter $\{I\}$ on I, prove that

$$\prod \mathfrak{A}_i/F \simeq \langle \prod A_i, S\rangle.$$

§ 2. Łoś's theorem

Let $\{\mathfrak{A}_i : i \in I\}$ be a collection of relational structures of the same type, and let F be an ultrafilter on I. It is clear from the construction we have described in the previous section that the ultraproduct $\prod \mathfrak{A}_i/F$ is of the same type as each of the $\mathfrak{A}_i$. Thus we can use the same language L to make statements about $\prod \mathfrak{A}_i/F$ as we use to talk about the structures from which it is built. The question naturally arises as to how the elementary properties of the ultraproduct are related to those of its constituent factors. Łoś's theorem, which we prove below, gives a complete answer to this question.

If $f = \langle f_1, \ldots, f_n, \ldots\rangle$ is a sequence of elements of $\prod A_i$, that is $f \in (\prod A_i)^\omega$, we let f/F be the sequence

$$f/F = \langle f_1/F, \ldots, f_n/F, \ldots\rangle$$

of elements of $\prod A_i/F$. Thus

$$f/F \in (\prod A_i/F)^\omega.$$

We let $f(i)$ be the sequence

$$f(i) = \langle f_1(i), \ldots, f_n(i), \ldots\rangle$$

of elements of A_i.

We are now ready to state the fundamental result on which the theory of ultraproducts depends.

THEOREM 2.1. ŁOŚ'S THEOREM

If F is an ultrafilter on I, for any formula ϕ of L *and any sequence* $f/F \in (\prod A_i/F)^{\omega}$

$$\prod \mathfrak{A}_i/F \vDash_{f/F} \phi \quad \textit{iff} \quad \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \phi\} \in F.$$

PROOF. The proof is by induction on the number of logical symbols in ϕ. Suppose first that ϕ is an atomic formula of the form $v_m = v_n$. Then

$$\begin{aligned} \prod \mathfrak{A}_i/F \vDash_{f/F} v_m = v_n \quad &\text{iff} \quad f_m/F = f_n/F \\ &\text{iff} \quad f_m \sim_F f_n \\ &\text{iff} \quad \{i \in I : f_m(i) = f_n(i)\} \in F \\ &\text{iff} \quad \{i \in I : \mathfrak{A}_i \vDash_{f(i)} v_m = v_n\} \in F. \end{aligned}$$

A similar argument shows that the theorem also holds if ϕ is an atomic formula of the form $P_0(v_m, v_n)$. Therefore it is true for all atomic formulas. This forms the basis of the induction.

Suppose now that the result is true for all formulas containing less than r logical symbols and that ϕ is a formula containing r logical symbols. There are three cases to consider:

(1). ϕ is of the form $\psi \wedge \chi$. Let

$$D_\psi = \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \psi\}$$

and

$$D_\chi = \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \chi\}.$$

Then

$$D_\psi \cap D_\chi = \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \psi \wedge \chi\}.$$

Now

$$\begin{aligned} \prod \mathfrak{A}_i/F \vDash_{f/F} \psi \wedge \chi \quad &\text{iff} \quad \prod \mathfrak{A}_i/F \vDash_{f/F} \psi \quad \text{and} \quad \prod \mathfrak{A}_i/F \vDash_{f/F} \chi, \\ &\text{iff} \quad D_\psi \text{ and } D_\chi \text{ are in } F, \text{ by the induction hypothesis}, \\ &\text{iff} \quad D_\psi \cap D_\chi \in F, \text{ since } F \text{ is a filter}. \end{aligned}$$

This shows that the theorem is true for $\psi \wedge \chi$.

(2). ϕ is of the form $\neg\psi$.

$$\begin{aligned} \prod \mathfrak{A}_i/F \vDash_{f/F} \neg\psi \quad &\text{iff} \quad \text{not} \prod \mathfrak{A}_i/F \vDash_{f/F} \psi, \\ &\text{iff} \quad \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \psi\} \notin F, \text{ by the induction hypothesis}, \\ &\text{iff} \quad I - \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \psi\} \in F, \text{ since } F \text{ is an ultrafilter}, \\ &\text{iff} \quad \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \neg\psi\} \in F. \end{aligned}$$

(Note that this is the only place in the proof where we use the fact that F is an ultrafilter rather than just a filter.)

(3). ϕ is of the form $(\exists v_n)\psi$. Let

$$D=\{i\in I:\mathfrak{A}_i\vDash_{f(i)}(\exists v_n)\psi\}.$$

Then we have to show that

$$\prod\mathfrak{A}_i/F\vDash_{f/F}(\exists v_n)\psi \quad \text{iff} \quad D\in F.$$

Suppose

$$\prod\mathfrak{A}_i/F\vDash_{f/F}(\exists v_n)\psi.$$

Then there is some $b\in\prod A_i$ such that

$$\prod\mathfrak{A}_i/F\vDash_{f(n/b)/F}\psi.$$

Let $E=\{i\in I:\mathfrak{A}_i\vDash_{f(n/b)(i)}\psi\}$. By the induction hypothesis $E\in F$.

$$f(n/b)(i)=f(i)(n/b_i)$$

and therefore $E\subseteq D$, and so because F is a filter, $D\in F$.

Conversely suppose that $D\in F$. If $i\in D$,

$$\mathfrak{A}_i\vDash_{f(i)}(\exists v_n)\psi$$

and hence there is some $b_i\in A_i$ such that

$$\mathfrak{A}_i\vDash_{f(i)(n/b_i)}\psi.$$

By the axiom of choice there is a $c\in\prod A_i$ such that $c_i=b_i$ for each $i\in D$, and is an arbitrary element of A_i otherwise. Then

$$D\subseteq\{i\in I:\mathfrak{A}_i\vDash_{f(n/c)(i)}\psi\}=C, \text{ say.}$$

Therefore $C\in F$ and hence by the induction hypothesis,

$$\prod\mathfrak{A}_i/F\vDash_{f(n/c)}\psi$$

and so

$$\prod\mathfrak{A}_i/F\vDash_{f/F}(\exists v_n)\psi.$$

This completes the proof of the theorem.

In particular, for sentences, we have:

COROLLARY 2.2. *If σ is a sentence of* L,

$$\prod\mathfrak{A}_i/F\vDash\sigma \quad \text{iff} \quad \{i\in I:\mathfrak{A}_i\vDash\sigma\}\in F.$$

PROOF. Obvious from theorem 2.1.

Intuitively this corollary says that a sentence holds in $\prod\mathfrak{A}_i/F$ iff it holds in 'almost all' the factors.

Let $\mathfrak{A}$ be a relational structure and let F be an ultrafilter on the set I. For each $a \in A$ we let $a^* \in A^I$ be such that $a^*(i) = a$, for all $i \in I$. We define the map $d: A \to A^I/F$, by

$$d(a) = a^*/F.$$

d is called the *canonical embedding* of $\mathfrak{A}$ into $\mathfrak{A}^I/F$. The following consequence of Łoš's theorem shows that in fact it is an elementary embedding.

LEMMA 2.3. *The canonical embedding is an elementary embedding of* $\mathfrak{A}$ *into* $\mathfrak{A}^I/F$, *and hence* $\mathfrak{A} \equiv \mathfrak{A}^I/F$.

PROOF. Let $\phi(v_0,\ldots,v_n)$ be any formula of L and let $a_0,\ldots,a_n$ be any elements of A. Then, by Łoš's theorem,

$$\begin{aligned} \mathfrak{A}^I/F \models \phi[a_0^*/F,\ldots,a_n^*/F] \quad &\text{iff} \quad \{i \in I : \mathfrak{A} \models \phi[a_0^*(i),\ldots,a_n^*(i)]\} \in F \\ &\text{iff} \quad \{i \in I : \mathfrak{A} \models \phi[a_0,\ldots,a_n]\} \in F \\ &\text{iff} \quad \mathfrak{A} \models \phi[a_0,\ldots,a_n]. \end{aligned}$$

Therefore d is an elementary embedding.

EXERCISE 2.4. Characterize those formulas of L for which Łoš's theorem holds when F is only a filter on I, and not an ultrafilter.

§ 3. Finite axiomatizability

Again we assume we are considering relational structures of some fixed type. A property P of these structures is called a *first order property* if there is some sentence σ such that

$$\mathfrak{A} \models \sigma \quad \text{iff} \quad \mathfrak{A} \text{ has the property } P.$$

Alternatively, if P is a first order property, we say that the notion of having the property P is *finitely axiomatizable.*

EXAMPLES

3.1. For each integer n let $\exists^{\geq n}$ be an abbreviation for the sentence

$$(\exists v_1)\cdots(\exists v_n)[v_1 \neq v_2 \wedge \cdots \wedge v_1 \neq v_n \wedge v_2 \neq v_3 \wedge \cdots \wedge v_{n-1} \neq v_n].$$

(Notice that we use $v_i \neq v_j$ as an abbreviation for $\neg v_i = v_j$.)

It can easily be seen that $\mathfrak{A} \models \exists^{\geq n}$ iff A contains at least n elements. So the property of containing at least n elements is a first order property.

3.2. Let $\exists!^n$ be an abbreviation for the sentence

$$\exists^{\geq n} \wedge \neg \exists^{\geq n+1},$$

then $\mathfrak{A} \models \exists!^n$ iff A contains precisely n elements. Thus the property of being of cardinal n is, for $n<\omega$, a first order property.

By a *general first order property* we mean a property P such that there is a set Σ of sentences such that a relational structure has the property iff it is a model of Σ.

EXAMPLE 3.3. Let $\Sigma=\{\exists^{\geq n}: n<\omega\}$. Clearly $\mathfrak{A}$ is infinite iff $\mathfrak{A}$ is a model of Σ. Thus the property of being infinite is a general first order property.

It is natural to inquire whether the property of being finite is a first order property, or equivalently, whether the property of being infinite is a first order property. A simple consequence of Łoš's theorem shows that the answer is no.

THEOREM 3.4. *Let Σ be a set of sentences. If there are arbitrarily large finite models of Σ, then there is an infinite model of Σ.*

PROOF. If Σ has arbitrarily large finite models, there is a sequence $\langle \mathfrak{A}_r : r<\omega\rangle$ of finite models of Σ such that if A_r is of cardinal n_r, then $n_r<n_s$ whenever $r<s<\omega$.

Let F be a non-principal ultrafilter on ω, and let $\mathfrak{A}$ be the ultraproduct $\prod \mathfrak{A}_r/F$. Since each sentence of Σ holds in each $\mathfrak{A}_r$ it is immediate from Łoš's theorem that $\mathfrak{A}$ is also a model of Σ. For each n, the sentence $\exists!^n$ holds in at most one $\mathfrak{A}_r$. F being non-principal contains no one element sets. Therefore, using Łoš's theorem again, $\exists!^n$ does not hold in $\mathfrak{A}$, for any $n<\omega$. Therefore $\mathfrak{A}$ is infinite.

COROLLARY 3.5. *The property of being a finite set is not a general first order property.*

PROOF. Since there are arbitrarily large finite sets but no infinite finite sets this is an immediate consequence of theorem 3.4.

Before we give another important consequence we show that by contrast, several other familiar notions are finitely axiomatizable.

EXAMPLE 3.6. Let σ_0 be the sentence

$$(\forall v_0)\neg P(v_0,v_0) \wedge (\forall v_0)(\forall v_1)[P(v_0,v_1) \vee P(v_1,v_0) \vee v_0=v_1] \wedge (\forall v_0)(\forall v_1)(\forall v_2)[P(v_0,v_1) \wedge P(v_1,v_2) \rightarrow P(v_0,v_2)],$$

then $\mathfrak{A}=\langle A,R\rangle$ is a model of σ_0 iff R is a (strict) total ordering of A.

EXAMPLE 3.7. Let σ_{OL} be the sentence

$$\sigma_{\mathrm{O}} \wedge (\exists v_0)(\forall v_1) \neg P(v_0,v_1),$$

then $\mathfrak{A}=\langle A,R\rangle$ is a model of σ_{OL} iff R is a total ordering of A with a largest element.

EXAMPLE 3.8. Let σ_{DOL} be the sentence

$$\sigma_{\mathrm{OL}} \wedge (\forall v_0)[(\exists v_1) P(v_1,v_0) \rightarrow (\exists v_2)[P(v_2,v_0) \wedge (\forall v_3)(\neg[P(v_2,v_3) \wedge P(v_3,v_0)])]],$$

then $\mathfrak{A}=\langle A,R\rangle$ is a model of σ_{DOL} iff R is a total ordering of A with a largest element in which each element with some element less than it has an immediate predecessor.

Let P_0 be a first order property and let P_1 be any property of relational structures. Clearly the property of 'having the properties P_0 and P_1' is a general first order property if P_1 is a general first order property. We use this in the proof of the following consequence of theorem 3.4.

THEOREM 3.9. *The property of being a well-ordered set is not a general first order property.*

PROOF. Let σ_{DOL} be the sentence of example 3.8. There are arbitrarily large finite well-ordered models of σ_{DOL} but no infinite well-ordered models of σ_{DOL}. The result now follows immediately from theorem 3.4 and the remarks above.

We give another proof of this result, which proves something even stronger, in ch. 7, lemma 1.13.

To each collection K of relational structures there corresponds the property of being an element of K. We describe necessary and sufficient algebraic conditions on K for this property to be a general first order property in chs. 7 and 8. We conclude this section with more examples of finitely axiomatizable notions.

EXAMPLE 3.10. DENSELY ORDERED SETS

Let σ_{DOS} be the sentence

$$\sigma_{\mathrm{O}} \wedge (\forall v_0)(\forall v_1)[P(v_0,v_1) \rightarrow (\exists v_2)[P(v_0,v_2) \wedge P(v_2,v_1)]].$$

Then $\mathfrak{A}$ is a densely ordered set iff $\mathfrak{A}$ is a model of σ_{DOS}.

EXAMPLE 3.11. GROUPS

In this case the appropriate language has one ternary predicate letter S (to represent the group addition) and one constant symbol e (whose value in any model will be the group unit). Let σ_G be the conjunction of the following sentences:

$\forall x \forall y \exists z \forall w [S(x,y,w) \leftrightarrow w = z]$ (existence and uniqueness of sum),

$\forall x \forall y \forall z \forall u \forall v \forall w [S(x,y,u) \wedge S(u,z,v) \wedge S(y,z,w) \rightarrow S(x,w,v)]$ (associativity of addition),

$\forall x S(x,e,x)$ (e is a right unit),

$\forall x \exists y S(x,y,e)$ (existence of right inverses),

then $\mathfrak{A}$ is a group iff $\mathfrak{A} \models \sigma_G$.

EXAMPLE 3.12. COMMUTATIVE GROUPS

Let σ_{CG} be the conjunction of σ_G and the sentence

$$\forall x \forall y \forall z [S(x,y,z) \rightarrow S(y,x,z)],$$

then $\mathfrak{A}$ is a commutative group iff $\mathfrak{A} \models \sigma_{CG}$.

EXAMPLE 3.13. RINGS

Here the appropriate language contains two ternary predicate letters S (to represent ring addition) and P (to represent ring multiplication), and one constant symbol 0 whose value will be the zero of the ring. Let σ'_{CG} be the result of substituting 0 for e in σ_{CG}, and let σ_R be the conjunction of σ'_{CG} and the following sentences:

$\forall x \forall y \exists z \forall w [P(x,y,w) \leftrightarrow w = z]$ (existence and uniqueness of product),

$\forall x \forall y \forall z \forall u \forall v \forall w [P(x,y,z) \wedge P(z,u,v) \wedge P(y,u,w) \rightarrow P(x,w,v)]$ (associativity of multiplication),

$\forall x \forall y \forall z \forall t \forall u \forall v \forall w [S(x,y,z) \wedge P(z,t,u) \wedge P(x,t,v) \wedge P(y,t,w) \rightarrow S(v,w,u)]$

$\forall x \forall y \forall z \forall t \forall u \forall v \forall w [S(x,y,z) \wedge P(t,z,u) \wedge P(t,x,v) \wedge P(t,y,w) \rightarrow S(v,w,u)]$ (distributivity of multiplication over addition),

then $\mathfrak{A} \models \sigma_R$ iff $\mathfrak{A}$ is a ring.

EXAMPLE 3.14. COMMUTATIVE RINGS

Let σ_{CR} be the conjunction of σ_R and the sentence

$$\forall x \forall y \forall z [P(x,y,z) \rightarrow P(y,x,z)],$$

then $\mathfrak{A} \models \sigma_{CR}$ iff $\mathfrak{A}$ is a commutative ring.

EXAMPLE 3.15. COMMUTATIVE FIELDS
Let σ_F be the conjunction of σ_{CR} and the sentences

$\exists x \exists y (\neg x = y)$ (existence of at least two elements),
$\forall x [\neg x = 0 \to \forall y \exists z P(x,z,y)]$ (solution of $x \cdot z = y$ if $x \neq 0$),
$\forall x [\neg x = 0 \to \forall y \exists z P(z,x,y)]$ (solution of $z \cdot x = y$ if $x \neq 0$),

then $\mathfrak{A}$ is a commutative field iff $\mathfrak{A}$ is a model of σ_F.

EXAMPLE 3.16. COMMUTATIVE FIELDS OF CHARACTERISTIC p
A field is said to be *of characteristic p* if p is the least integer such that, for each element x of the field,

$$p \cdot x = 0$$

where

$$p \cdot x = \underbrace{x + x + \cdots + x}_{p}.$$

To simplify the axioms for a field of characteristic p we use the following abbreviations.

$$S_1(x,y) \text{ abbreviates } y = x$$

and

$$S_{n+1}(x,y) \text{ abbreviates } (\exists z)[S_n(x,z) \wedge S(z,x,y)]$$

thus $S_n(x,y)$ expresses the fact that $y = n \cdot x$.

Let σ_F^p be the conjunction of σ_{CF} and the sentence

$$C_p = (\forall x) S_p(x,0) \wedge \neg [(\forall x) S_{p-1}(x,0) \vee (\forall x) S_{p-2}(x,0) \vee \ldots \vee (\forall x) S_1(x,0)]$$

Then $\mathfrak{A}$ is a commutative field of characteristic p iff $\mathfrak{A}$ is a model of σ_F^p. It is well known that there is a field of characteristic p iff p is a prime number. Also if $\mathfrak{A}$ is a field of characteristic p and $q \neq p$, then $\mathfrak{A}$ is not a field of characteristic q.

A field of *characteristic zero* is a field which is not of characteristic p for any prime p (for example, the field of rational numbers). Let

$$\Delta_0 = \{\sigma_F\} \cup \{\neg C_p : p < \omega\},$$

where C_p is the sentence defined above. $\mathfrak{A}$ is a field of characteristic zero iff $\mathfrak{A}$ is a model of Δ_0. This shows that the property of being a field of characteristic zero is a general first order property.

The next theorem shows that Δ_0 cannot be replaced by any finite set of sentences.

THEOREM 3.17. *The property of being a (commutative) field of characteristic zero is not a first order property.*

PROOF. Suppose there is a sentence σ_F^0 such that $\mathfrak{A} \models \sigma_F^0$ iff $\mathfrak{A}$ is a commutative field of characteristic zero.

For each prime p let $\mathfrak{A}_p$ be a commutative field of characteristic p, that is, a model of $\{\sigma_F, C_p\}$. Let F be a non-principal ultrafilter on the set of all primes and let

$$\mathfrak{A} = \prod_p \mathfrak{A}_p / F .$$

Since each $\mathfrak{A}_p$ is a model of σ_F, by Łoš's theorem, $\mathfrak{A}$ is also a model of σ_F. For each p, C_p holds in precisely one of the factors $\mathfrak{A}_q$ (the one with $q = p$). F is a non-principal ultrafilter and therefore contains no one element sets. Hence, using Łoš's theorem again, $\mathfrak{A}$ is not a model of C_p. This shows that $\mathfrak{A}$ is of characteristic zero and so $\mathfrak{A} \models \sigma_F^0$. Since σ_F^0 does not hold in any of the factors $\mathfrak{A}_p$, this contradicts Łoš's theorem as the empty set is not in F. This completes the proof.

EXAMPLE 3.18. VECTOR SPACES

A vector space as ordinarily conceived does not consist of a single set with a number of relations defined on it but of two disjoint sets – a field of scalars and a commutative group of vectors – together with various relations between them. So it might seem that the notion of a vector space cannot be expressed in a first order predicate language where all the variables range over the same set. However this difficulty can be overcome and we describe one way of doing this here.

The language L that we use has two unary predicate letters

$$V, F,$$

four ternary predicate letters

$$S, P, A, M,$$

and two constant symbols

$$0, \mathbf{0}.$$

Intuitively, $V(x)$ says 'x is a vector', $F(x)$ says 'x is a scalar', we use S and P to represent the addition and multiplication of scalars, A to represent the addition of vectors and M to represent the multiplication of a vector by a scalar. 0 represents the zero scalar and $\mathbf{0}$ the zero vector.

We let σ_{VS} be the conjunction of the following sentences of L:

(1). $(\forall x)[V(x) \leftrightarrow \neg F(x)]$,

'Everything is either a vector or a scalar, but not both.'

(2). $(\forall u)(\forall v)(\forall w)([S(u,v,w) \vee P(u,v,w)] \rightarrow [F(u) \wedge F(v) \wedge F(w)])$,
'S and P are relations between scalars.'

(3). $(\forall x)(\forall y)(\forall z)(A(x,y,z) \rightarrow [V(x) \wedge V(y) \wedge V(z)])$,
'A is a relation between vectors.'

(4). $(\forall u)(\forall x)(\forall y)(M(u,x,y) \rightarrow [F(u) \wedge V(x) \wedge V(y)])$,
'M is a relation between one scalar and two vectors.'

(5). The sentence which says that the scalars form a commutative field with addition S, multiplication P and zero 0. This sentence can be derived from the sentence σ_F of example 3.15 in an obvious way.

(6). The sentence which says that the vectors form a commutative group with addition A and zero $\mathbf{0}$. This sentence can be derived from the sentence σ_{CG} of example 3.12 in an obvious way.

(7). Sentences which say how scalars operate on vectors. We can take these to be the following:

$(\forall u)(\forall x)([F(u) \wedge V(x)] \rightarrow (\exists y)(\forall z)[M(u,x,z) \leftrightarrow z = y])$,
'Given any scalar u and any vector x, there is a unique vector $u \cdot x$.'

$(\forall u)(\forall v)(\forall w)(\forall x)(\forall y)(\forall z)(\forall t)([S(u,v,w) \wedge M(u,x,y) \wedge M(v,x,z) \wedge A(y,z,t)] \rightarrow M(w,x,t))$,
'For any scalars u,v and any vector x, $(u+v) \cdot x = u \cdot x + v \cdot x$.'

$(\forall u)(\forall x)(\forall y)(\forall z)(\forall x')(\forall y')(\forall z')([A(x,y,z) \wedge M(u,x,x') \wedge M(u,y,y') \wedge A(x',y',z')] \rightarrow M(u,z,z'))$,
'For any scalar u and any vectors x, y, $u \cdot (x+y) = u \cdot x + u \cdot y$.'

$(\forall u)(\forall v)(\forall w)(\forall x)(\forall y)(\forall z)([P(u,v,w) \wedge M(w,x,y) \wedge M(v,x,z)] \rightarrow M(u,z,y))$,
'For any scalars u,v and any vector x, $(uv) \cdot x = u \cdot (v \cdot x)$.'

$(\forall u)[(\forall v) P(u,v,v) \rightarrow (\forall x) M(u,x,x)]$,
'For any vector x, $1 \cdot x = x$.'

Then $\mathfrak{A} = \langle A, F_{\mathfrak{A}}, V_{\mathfrak{A}}, S_{\mathfrak{A}}, P_{\mathfrak{A}}, A_{\mathfrak{A}}, M_{\mathfrak{A}}, 0_{\mathfrak{A}}, \mathbf{0}_{\mathfrak{A}} \rangle$ is a model of σ_{VS} iff $F_{\mathfrak{A}}$ and $V_{\mathfrak{A}}$ are disjoint sets whose union is A, and $\langle V_{\mathfrak{A}}, A_{\mathfrak{A}}, \mathbf{0}_{\mathfrak{A}} \rangle$ is a vector space over the field $\langle F_{\mathfrak{A}}, S_{\mathfrak{A}}, P_{\mathfrak{A}}, 0_{\mathfrak{A}} \rangle$ with $M_{\mathfrak{A}}$ as the multiplication of a vector by a scalar. Thus we can regard the property of being a vector space as a first order property.

EXAMPLE 3.19. FINITE DIMENSIONAL VECTOR SPACES

A vector space is of dimension n if there are n linearly independent vectors which span the space. We can express this property in the language L

described in the previous example. For convenience we adopt the following abbreviations:

$S_n(u_1,\ldots,u_n;x_1,\ldots,x_n;y)$ is an abbreviation for

$$(\exists z_1)\ldots(\exists z_n)(\exists t_1)\ldots(\exists t_{n-2})[M(u_1,x_1,z_1)\wedge\cdots\wedge M(u_n,x_n,z_n)$$
$$\wedge A(z_1,z_2,t_1)\wedge A(t_1,z_3,t_2)\wedge\cdots\wedge A(t_{n-3},z_{n-1},t_{n-2})\wedge A(t_{n-2},z_n,y)].$$

Thus $S_n(u_1,\ldots,u_n;x_1,\ldots,x_n;y)$ 'says' that

$$y=u_1\cdot x_1+\cdots+u_n\cdot x_n.$$

Also we let $LI_n(x_1,\ldots,x_n)$ be an abbreviation for

$$(\forall u_1)\ldots(\forall u_n)[S_n(u_1,\ldots,u_n;x_1,\ldots,x_n;\mathbf{0})\rightarrow(u_1=0\wedge\cdots\wedge u_n=0)].$$

$LI_n(x_1,\ldots,x_n)$ 'says' that the vectors $\{x_1,\ldots,x_n\}$ form a linearly independent set.

Now we let $\tau_{\mathrm{Dim}(n)}$ be the sentence

$$(\exists x_1)\ldots(\exists x_n)(LI_n(x_1,\ldots,x_n)$$
$$\wedge(\forall y)[V(y)\rightarrow(\exists u_1)\ldots(\exists u_n)S_n(u_1,\ldots,u_n;x_1,\ldots,x_n;y)]).$$

Then a vector space is of dimension n iff it is a model of $\tau_{\mathrm{Dim}(n)}$.

Let $\Sigma=\{\sigma_{\mathrm{VS}}\}\cup\{\neg\tau_{\mathrm{Dim}(n)}:n\in\omega\}$. Then $\mathfrak{A}$ is an infinite dimensional vector space iff it is a model of Σ. An argument similar to that used to prove theorem 3.17 shows that Σ cannot be replaced by a finite set of sentences. Thus the notion of an infinite dimensional vector space is a general first order property but not a first order property. It follows that the property of being a finite dimensional vector space is not a first order property.

For a further discussion of this see example 7.1.10.

EXAMPLE 3.20. ALGEBRAICALLY CLOSED FIELDS

A field is said to be *algebraically closed* if each non-constant polynomial with coefficients in the field has a zero in the field, or, equivalently, if each such polynomial factors into linear factors over the field. In order to write down axioms for algebraically closed fields it is convenient to introduce the following abbreviations. We shall be using the language for fields introduced in example 3.15.

Let $M_0(x,y)$ be an abbreviation for $y=x$ and let $M_{k+1}(x,y)$ be an abbreviation for the formula

$$(\exists z)[M_k(x,z)\wedge P(z,x,y)],$$

thus $M_k(x,y)$ 'says' $y=x^{k+1}$.

Next for $n \geqslant 2$, we let $T_n(x_0,\ldots,x_n,y,z)$ be an abbreviation for

$$(\exists u_0)\ldots(\exists u_n)(\exists v_0)\ldots(\exists v_n)(\exists w_1)\ldots(\exists w_{n-1})$$
$$[M_0(y,u_0)\wedge\cdots\wedge M_n(y,u_n)\wedge P(u_0,x_0,v_0)\wedge\cdots\wedge P(u_n,x_n,v_n)$$
$$\wedge S(v_0,v_1,w_1)\wedge S(w_1,v_2,w_2)\wedge\cdots\wedge S(w_{n-1},v_n,z)],$$

thus $T_n(x_0,x_1,\ldots,x_n,y,z)$ 'says' $z=x_0y^0+x_1y^1+\cdots+x_ny^n$.

Finally for $n\geqslant 2$, we let τ_n be the sentence

$$(\forall x_0)\ldots(\forall x_n)(\neg x_n=0\rightarrow(\exists y)\,T_n(x_0,\ldots,x_n,y,0)).$$

τ_n is the sentence which says that all polynomials of degree n have a zero in the field.

Then if

$$\Sigma_{\mathrm{CF}}=\{\sigma_{\mathrm{F}}\}\cup\{\tau_n:2\leqslant n<\omega\},$$

$\mathfrak{A}$ is an algebraically closed field iff it is a model of Σ_{CF}. Thus the property of being an algebraically closed field is a general first order property.

The next algebraic results enables us to deduce that the theory of algebraically closed fields is not finitely axiomatizable.

THEOREM 3.21. *For each $n\in\omega$, there is a field $\mathfrak{A}_n$, which is not algebraically closed, in which all non-constant polynomials of degree $\leqslant n$ have zeros.*

PROOF. Let $\boldsymbol{Q}$ be the field of rational numbers and let $\overline{\boldsymbol{Q}}$ be its algebraic closure. $\overline{\boldsymbol{Q}}$ is countable and so can be enumerated in a sequence $\langle q_i:i\in\omega\rangle$.

We define the sequence $\langle \boldsymbol{Q}^n_r: r\in\omega\rangle$ of fields as follows.

$$\boldsymbol{Q}^n_0=\boldsymbol{Q}$$

and

$$\boldsymbol{Q}^n_{r+1}=\boldsymbol{Q}^n_r[q],$$

where q is the first element in the given enumeration of $\overline{\boldsymbol{Q}}$ which is not in $\boldsymbol{Q}^n_r$ and which is a zero of a polynomial of degree $\leqslant n$, which has coefficients in $\boldsymbol{Q}^n_r$.

Now we let $\mathfrak{A}_n=\bigcup_{r\in\omega}\boldsymbol{Q}^n_r$. It is easy to see that $\mathfrak{A}_n$ is a field. Suppose $f(x)=a_tx^t+\cdots+a_1x+a_0$ is a polynomial of degree $\leqslant n$, with coefficients in $\mathfrak{A}_n$. For some $r\in\omega$, $a_0,\ldots,a_t$ are all in $\boldsymbol{Q}^n_r$. $f(x)$ has a zero in $\overline{\boldsymbol{Q}}$, say q_i in the given enumeration. Then it is clear from our construction that $q_i\in\boldsymbol{Q}^n_{r+i}$ and so $f(x)$ has a zero in $\mathfrak{A}_n$.

We complete the proof by showing that $\mathfrak{A}_n$ is not algebraically closed. Let p be any prime greater than n, and for each $r\in\omega$, let $n_r=[\boldsymbol{Q}^n_{r+1}:\boldsymbol{Q}^n_r]$ be

the degree of $\mathbb{Q}_{r+1}^n$ over $\mathbb{Q}_r^n$. By construction $n_r \leqslant n$ and so p does not divide n_r.

By a very well known result, for each $s \in \omega$

$$[\mathbb{Q}_s^n : \mathbb{Q}] = [\mathbb{Q}_s^n : \mathbb{Q}_{s-1}^n] \times [\mathbb{Q}_{s-1}^n : \mathbb{Q}_{s-2}^n] \times \cdots \times [\mathbb{Q}_1^n : \mathbb{Q}_0^n]$$
$$= n_{s-1} \cdot n_{s-2} \cdot \cdots \cdot n_0,$$

and so $p \not| [\mathbb{Q}_s^n : \mathbb{Q}]$.

It follows from this that none of the elements of $\mathfrak{A}_n$ have degrees over $\mathbb{Q}$ which are divisible by p. But $\bar{\mathbb{Q}}$ contains elements which are of degree p over $\mathbb{Q}$, for example all the zeros of $x^p - 2$. This shows that $\mathfrak{A}_n$ is not algebraically closed.

THEOREM 3.22. *The theory of algebraically closed fields is not finitely axiomatizable.*

PROOF. Suppose there is a sentence σ such that a structure is an algebraically closed field iff it is a model of σ.

For each n let $\mathfrak{A}_n$ be the field given by the previous theorem and let F be a non-principal ultrafilter on ω. Let $\mathfrak{A} = \prod \mathfrak{A}_n / F$.

Each $\mathfrak{A}_n$ is a model of σ_F and so, by Łoš's theorem, so too is $\mathfrak{A}$, thus $\mathfrak{A}$ is certainly a field. Also each of the sentences τ_n holds in all but a finite number of the factors of $\mathfrak{A}$, therefore again using Łoš's theorem and the fact that F is a non-principal ultrafilter, we can conclude that each sentence τ_n holds in $\mathfrak{A}$. Therefore $\mathfrak{A}$ is an algebraically closed field, and so, by hypothesis, is a model of σ. But each $\mathfrak{A}_n$ is not algebraically closed and so is a model of $\neg\sigma$. Therefore, by Łoš's theorem, $\mathfrak{A}$ is also a model of $\neg\sigma$. This contradiction shows that no such sentence σ exists.

COROLLARY 3.23 *The theory of algebraically closed fields of characteristic zero is not finitely axiomatizable.*

PROOF. It is sufficient to observe that all the fields $\mathfrak{A}_n$ are of characteristic zero.

COROLLARY 3.24. *For each prime p, the theory of algebraically closed fields of characteristic p is not finitely axiomatizable.*

PROOF. This is proved similarly to the previous result. It is necessary only to repeat the construction of theorem 3.21 starting not with $\mathbb{Q}$ but with the prime field of characteristic p, that is, the field of integers modulo p. In this case all the fields $\mathfrak{A}_n$ will be of characteristic p and so the argument is the same as before.

§4. The compactness theorem

It can be seen that the arguments we used to prove theorems 3.4 and 3.17 did not involve the nature of the sentences $\exists^{\geqslant n}$ and C_p in any essential way and the argument was basically the same in each case. This same basic argument can be used to prove the following powerful general result.

THEOREM 4.1. THE COMPACTNESS THEOREM FOR PREDICATE CALCULUS
A set Σ of sentences of L *has a model iff each finite subset of Σ has a model.*

PROOF. If Σ has a model then, trivially, each finite subset of Σ has a model.

Suppose that each finite subset of Σ has a model. We are going to construct an ultraproduct of the models of the finite subsets of Σ. By choosing the ultrafilter carefully we ensure that the ultraproduct we construct is a model of the whole set Σ.

Let $I = S_\omega(\Sigma)$. By hypothesis, for each $\Delta \in I$ there is a model $\mathfrak{A}_\Delta$ of Δ.

For each set Δ in I let $\Delta^* = \{\Delta' \in I : \Delta \subseteq \Delta'\}$. Each collection Δ^* is non-empty. Also for any finite subset, say $\{\Delta_1, \ldots, \Delta_n\}$ of I,

$$\Delta_1 \cup \cdots \cup \Delta_n \in \Delta_1^* \cap \cdots \cap \Delta_n^*$$

and hence the collection $\{\Delta^* : \Delta \in I\}$ has the finite intersection property. Therefore this collection can be extended to an ultrafilter F on I. We show that the ultraproduct

$$\prod_{\Delta \in I} \mathfrak{A}_\Delta / F$$

is a model of Σ.

Suppose $\sigma \in \Sigma$, then $\{\sigma\} \in I$, say $\{\sigma\} = \Delta_0$. $\mathfrak{A}_{\Delta_0} \models \sigma$ and, clearly, $\mathfrak{A}_{\Delta'} \models \sigma$, if $\Delta_0 \subseteq \Delta'$. Hence

$$\Delta_0^* = \{\Delta' \in I : \Delta_0 \subseteq \Delta'\} \subseteq \{\Delta' \in I : \mathfrak{A}_{\Delta'} \models \sigma\}.$$

Since $\Delta_0^* \in F$, $\{\Delta' \in I : \mathfrak{A}_{\Delta'} \models \sigma\} \in F$ and therefore, by Łoš's theorem, $\prod \mathfrak{A}_\Delta / F \models \sigma$. This completes the proof.

The compactness theorem enables us to generalize corollary 3.5.3 to arbitrary sets of sentences.

THEOREM 4.2. THE GÖDEL–HENKIN COMPLETENESS THEOREM
A set Σ of sentences has a model iff it is consistent.

PROOF. By the finiteness theorem for predicate calculus, Σ is consistent iff each finite subset of Σ is consistent. By corollary 3.5.3, each finite subset of Σ is consistent iff each finite subset of Σ has a model. The compactness theorem shows that this holds iff Σ has a model.

Note that nowhere in theorems 4.1 and 4.2 we have used the fact that L is a countable language. Thus these theorems hold for languages with arbitrarily large sets of predicate letters.

The Gödel–Henkin completeness theorem together with the Löwenheim–Skolem theorem yields

THEOREM 4.3. *If Σ is a consistent set of sentences of cardinal α, then either Σ has a finite model or Σ has models of all infinite cardinals $\geqslant \alpha$.*

§5. The completeness theorem, the axiom of choice and the ultrafilter theorem

Our proofs of the compactness and completeness theorems involve the axiom of choice more or less explicitly. For instance, an assumption underlying the proof of the compactness theorem is that the ultraproduct constructed during this proof is *non-empty* and this can only be justified in general by appealing to the axiom of choice.

In this section we show that the strong form of the completeness theorem in which the cardinalities of the models are mentioned (theorem 4.3) is actually equivalent to the axiom of choice. In contrast, however, the compactness and the Gödel–Henkin completeness theorem (theorems 4.1 and 4.2) in which the cardinalities are not mentioned are equivalent to the ultrafilter theorem for Boolean algebras (theorem 1.3.4) which is known to be strictly weaker than the axiom of choice*. Our proofs of these results will need some facts about predicate calculus we have not so far mentioned; we give references to other textbooks where these results are explained.

THEOREM 5.1. *The strong completeness theorem is equivalent to the axiom of choice.*

PROOF. We have already remarked that the strong form of the completeness theorem can be derived from the axiom of choice.

Conversely, assume the result of theorem 4.3. Let σ be the sentence of theorem 4.3.7. Recall that σ has a model of cardinal α iff $\alpha^2=\alpha$. Let $\Sigma=\{\sigma,\exists^{\geqslant 2}\}$. Since $\aleph_0{}^2=\aleph_0$, Σ has a model of cardinal $\aleph_0$ and is therefore consistent. Clearly Σ has no finite models and therefore, by the strong form of the completeness theorem, has models of all cardinals $\alpha\geqslant\aleph_0$. Therefore for all infinite cardinals α, $\alpha^2=\alpha$ and we have already remarked that this is equivalent to the axiom of choice.

---

* This result was first proved by Halpern, Läuchli and Lévy (see the invaluable survey by MATHIAS [1967]).

To complete this section we show that the Gödel–Henkin completeness theorem, the compactness theorem and the ultrafilter theorem are all equivalent. Of course we must be careful to avoid any (tacit) applications of the axiom of choice in the proof.

THEOREM 5.2. *The following assertions are equivalent:*

(1). *Every Boolean algebra contains an ultrafilter.*

(2). *Each filter in a Boolean algebra can be extended to an ultrafilter.*

(3). *Each consistent set of sentences of first order predicate calculus has a model.*

(4). *If Σ is a set of first order sentences and every finite subset of Σ has a model then Σ itself has a model.*

PROOF. (1)$\Rightarrow$(2). Assume (1) and let F be a filter in a Boolean algebra B. Then the quotient algebra B/F is a Boolean algebra which according to (1) contains an ultrafilter U. Let h be the canonical homomorphism of B into B/F. It is easily checked that $h^{-1}[U]$ is an ultrafilter in B which extends F.

(2)$\Rightarrow$(3). We assume here a knowledge of Skolem transforms, which are explained for example in CHURCH [1956].

Assume (2) and let Σ be a consistent set of sentences, which we may take to be in prenex normal form, of a first order language L. For each formula ϕ of L in prenex normal form let ϕ^* be the *Skolem transform* of ϕ, that is, the formula obtained by taking the closure of ϕ, introducing Skolem functions and individual constants to eliminate existential quantifiers, and then suppressing the remaining universal quantifiers. (For example, the Skolem transform of the prenex formula $(\exists x)(\forall y)(\exists z)\phi(x,y,z)$ is $\phi(c,y,f(y))$, where c is a new individual constant and f a new unary function letter.) We can assume that the new function letters and individual constants introduced for one formula are distinct from those introduced for another formula. Let L* be a language whose set of primitive symbols includes all those of L, except the quantifiers, and also all the new function letters and individual constants we have introduced in forming the Skolem transforms of the formulas of L. The rules of inference for L* are to be modus ponens and substitution of terms for variables. As axioms for L* we take all instances of the schemas SC1, SC2 and SC3 of ch. 2, § 1, and also all instances of the schemas

$$\tau = \tau$$

and

$$\tau = \tau' \rightarrow [\phi(\tau) \rightarrow \phi(\tau')],$$

where τ, τ' are any terms and ϕ is any formula.

It can be shown, without making any use of the axiom of choice – and this is the crux of the proof – that Σ is consistent in L iff $\Sigma^* = \{\phi^* : \phi \in \Sigma\}$ is consistent in L*. (For a proof of this, the so-called *second ε-theorem*, consult HILBERT and BERNAYS [1939]).

Let B be the Lindenbaum algebra of L*, that is the Boolean algebra of equivalence classes of formulas of L* under the relation of provable equivalence in L*. For each formula ϕ of L* let $|\phi|$ be the equivalence class of ϕ under this relation, thus $|\phi| = \{\psi : \vdash_{L^*} \phi \leftrightarrow \psi\}$. Since Σ is consistent, so is Σ^* and hence the set $|\Sigma^*| = \{|\phi^*| : \phi^* \in \Sigma^*\}$ has the finite intersection property and so by lemma 1.2.7 can be extended to a filter in B and so, by (2), to an ultrafilter F in B. We now use the ultrafilter F to define an interpretation $\mathfrak{A}$ of L* as follows. The universe of $\mathfrak{A}$ is to be the set of all terms of L*, each constant is its own interpretation, and if f is an n-ary function letter we let its interpretation be the function F on the set of terms given by $F(\tau_1,\ldots,\tau_n) = f(\tau_1,\ldots,\tau_n)$. For each n-ary predicate letter P we stipulate that $P[\tau_1,\ldots,\tau_n]$ holds in $\mathfrak{A}$ iff $|P(\tau_1,\ldots,\tau_n)| \in F$. (For the purposes of this argument we regard the equality symbol '=' as a non-logical predicate letter, so it will not necessarily be interpreted as the identity relation on $\mathfrak{A}$.) Then it can easily be shown by induction that a formula ϕ of L* is valid in $\mathfrak{A}$ iff $|\phi| \in F$. But since $|\Sigma^*| \subseteq F$ it follows that every member of Σ^* is valid in $\mathfrak{A}$. A moment's reflection on how Σ^* was obtained from Σ will be enough to convince one that every member of Σ is also valid in $\mathfrak{A}$. Thus $\mathfrak{A}$ would be a model of Σ if the equivalence relation corresponding to the equality symbol were the identity relation. In general this will not be the case. If it is not we 'contract' $\mathfrak{A}$ to a model $\mathfrak{A}'$ which does satisfy this condition in the standard way (see, for example, MENDELSON [1964] p. 80). $\mathfrak{A}'$ is then a model of Σ, and (3) follows.

(3) ⇒ (4). Assume (3), and suppose that Σ is a set of first order sentences every finite subset of which has a model. Then clearly each finite subset of Σ is consistent and so Σ itself is consistent and so by (3) has a model.

(4) ⇒ (1). Assume (4), and suppose that $\mathfrak{B}$ is a Boolean algebra. Let L be the appropriate language for Boolean algebras and let L($\mathfrak{B}$) be the language obtained from L by adding a constant for each element of $\mathfrak{B}$. Let Σ_1 be the set of all open sentences of L($\mathfrak{B}$) which hold in $\mathfrak{B}$ when the constants are interpreted as the corresponding elements of $\mathfrak{B}$, and let Σ_2 be a set of axioms for Boolean algebras. Let L^+ be the language obtained from L($\mathfrak{B}$) by adding a new unary predicate letter U, and let Σ_3 be the set of sentences which say that the elements satisfying U form an ultrafilter. We will show that each finite subset of $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ has a model.

Let Σ_0 be a finite subset of Σ. Let A_0 be the set of those elements of $\mathfrak{B}$ corresponding to the constants occurring in Σ_0, and let $\mathfrak{A}$ be the finite subalgebra of $\mathfrak{B}$ generated by A_0 (i.e. the smallest subalgebra of $\mathfrak{B}$ containing all the elements in A_0). Since $\mathfrak{A}$ is finite we know, without using the axiom of choice, that $\mathfrak{A}$ contains an ultrafilter, say F. Thus if a is an enumeration of $\mathfrak{A}_0$, $(\mathfrak{A},F,a)$ is a model of Σ_0. Hence, by (4), Σ has a model $\mathfrak{B}'$ in which the value $U_{\mathfrak{B}'}$ of U is an ultrafilter. Since $\mathfrak{B}'$ is a model of Σ_1 it is, up to isomorphism, an extension of $\mathfrak{B}$. Thus $U_{\mathfrak{B}'} \cap B$, is an ultrafilter in $\mathfrak{B}$. This completes the proof.

§6. Historical and bibliographical remarks

The ultraproduct construction is foreshadowed by some earlier constructions of GÖDEL [1930] and SKOLEM [1934]. We discuss the relation between this early work and the ultraproduct construction in ch. 12. The construction was also used by HEWITT [1948] in connection with real closed fields. Its first explicit appearance in connection with model theory can be found in ŁOŚ [1955]. Łoś's theorem is implicit in this last paper.

Most of our remarks about finite axiomatizability are well known. TARSKI [1949] announced the result that the theory of algebraically closed fields is not finitely axiomatizable but as far as we know no proof has previously been published. Our proof of theorem 3.21 was inspired by a suggestion of A. Macintyre and H. Simmons who also suggested the extension of the result to algebraically closed fields of specified characteristic.

The compactness theorem for countable sets of sentences is due to GÖDEL [1930]. It was extended to uncountable languages by MALCEV [1936]. The proof using ultraproducts is due to MOREL, SCOTT and TARSKI [1958].

CHAPTER 6

SET THEORETIC PROPERTIES OF ULTRAPRODUCTS

Before going on to give some more applications of ultraproducts to model theory we discuss some of the purely set theoretic aspects of this construction.

§1. More about ultrafilters

Recall that if I is a set, by an *ultrafilter on* I we mean an ultrafilter in the power set Boolean algebra $\langle \mathscr{P}(I), \subseteq \rangle$. By an *ultrafilter pair* is meant a pair $\langle I,F \rangle$ where I is a non-empty set and F is an ultrafilter on I. In this section we are going to classify various sorts of ultrafilter pairs. We begin by repeating some of the results of ch. 1.

(a). *Principal ultrafilters.* A *principal* filter on I is a filter F of the form

$$F = \{Y \subseteq I : X \subseteq Y\},$$

where X is some fixed non-empty subset of I. We say that X *generates* the filter F. A *principal ultrafilter* is an ultrafilter which is a principal filter.

LEMMA 1.1. *If F is a filter on I, F is a principal filter iff $\bigcap \{X : X \in F\} \in F$.*

PROOF. If F is the principal filter generated by the set X_0, then $\bigcap \{X : X \in F\} = X_0 \in F$; while if $X_0 = \bigcap \{X : X \in F\}$ is in F, then F is the principal filter generated by X_0.

LEMMA 1.2. *If F is a principal filter, then F is an ultrafilter iff the set which generates F contains just one element.*

PROOF. Suppose $x \in I$, and $F = \{X \subseteq I : x \in X\}$. For each $X \subseteq I$ either $x \in X$ or $x \in I - X$, therefore by lemma 1.3.1, F is an ultrafilter.

Conversely suppose F is a principal ultrafilter and $X_0 = \bigcap \{X : X \in F\}$ contains the two elements x,y. Since F is an ultrafilter either $\{x\}$ or $I - \{x\}$ is in F. In the first case $y \notin X_0$ and in the second case $x \notin X_0$. This contradiction shows that X_0 contains just one element.

This lemma shows that if I is of cardinal α there are precisely α distinct principal ultrafilters on I. We have already noted that if I is finite all ultrafilters on I are principal. We show in the next section that the ultraproduct construction only provides something new when the ultrafilter is not principal, so from now on we shall only be interested in

(b). *Non-principal ultrafilters.* An ultrafilter is *non-principal* if it is not principal.

LEMMA 1.3. *A non-principal ultrafilter contains no finite sets.*

PROOF. Let F be an ultrafilter containing a finite set. Let X be a set of least cardinal in F. Since F is proper, X is not empty. We show that X contains just one element. Suppose not, say x and y are in X. By hypothesis $\{x\} \notin F$ and so $I-\{x\} \in F$. Therefore $X \cap (I-\{x\}) = X-\{x\} \in F$. But $X-\{x\}$ contains fewer elements than X since X is finite, contradicting our choice of X. This shows that X contains just one element, say x. It follows at once that F is the principal ultrafilter generated by $\{x\}$.

The existence of non-principal ultrafilters on infinite sets is guaranteed by the next result.

LEMMA 1.4. *If I is infinite, there is a non-principal ultrafilter on I.*

PROOF. Let $\bar{S}_\omega(I)$ be the set of all cofinite subsets of I. It is easily checked that $\bar{S}_\omega(I)$ has the finite intersection property and so can be extended to an ultrafilter F on I. By construction F contains no finite sets and therefore is non-principal.

Note that by lemma 1.3 any non-principal ultrafilter on I contains all the sets in $\bar{S}_\omega(I)$.

Another consequence of lemma 1.3 should also be noted. Suppose F is a non-principal ultrafilter on I and X, Y are subsets of I such that $X-Y$ is finite and $X \in F$. Then we can conclude that Y is also in F, because, since $X-Y$ is finite $I-(X-Y) \in F$, therefore $X \cap \{I-(X-Y)\} = X \cap Y \in F$. But $X \cap Y \subseteq Y$ and so $Y \in F$.

We have just shown that there is at least one non-principal ultrafilter on an infinite set I. Suppose I is of cardinal α. Then $\mathscr{P}(I)$ is of cardinal 2^α. An ultrafilter on I is a subset of $\mathscr{P}(I)$. There are 2^{2^α} subsets of $\mathscr{P}(I)$ and hence there are at most 2^{2^α} non-principal ultrafilters on I. The next theorem, due to Tarski, shows that this maximum number is actually attained.

THEOREM 1.5. *If I is an infinite set of cardinal α, there are 2^{2^α} non-principal ultrafilters on I.*

PROOF. Recall that $S_\omega(X)$ is the set of all finite subsets of X. Since I is infinite $S_\omega(I)$ also has cardinal α and $S_\omega(S_\omega(I))$ is also of cardinal α. In effect we show that there are 2^{2^α} ultrafilters on $S_\omega(S_\omega(I))$. The result then immediately generalizes to any set of cardinal α.

Since I is infinite there are two disjoint sets, A,B each of cardinal α, whose union is I. Let f be any one–one map of A onto B. If $X \subseteq A$ we let

$$X^+ = X \cup (B - \{f(x): x \in X\}),$$

thus X^+ is X together with those elements of B not in the image of X under f. Suppose $X, Y \subseteq A$ and $X \neq Y$. Then either $X - Y$ or $Y - X$ is non-empty. If $x \in X - Y$, $x \in X^+ - Y^+$, and if $x \in Y - X$, $f(x) \in X^+ - Y^+$. So in either case $X^+ \nsubseteq Y^+$. We have therefore shown that

$$\text{if } X, Y \subseteq A \text{ and } X \neq Y, \text{ then } X^+ \nsubseteq Y^+. \tag{1}$$

Let

$$\boldsymbol{I} = S_\omega(I).$$

$\boldsymbol{I}$ is also of cardinal α, and let

$$\mathscr{A} = \{\boldsymbol{I} - S_\omega(X^+): X \subseteq A\}.$$

Then $\mathscr{A} \subseteq \mathscr{P}(\boldsymbol{I})$ and by (1), $\mathscr{A}$ is of cardinal 2^α.

Suppose

$$\boldsymbol{X} = \boldsymbol{I} - S_\omega(X^+) \in \mathscr{A}$$

and $\mathscr{Y}$ is a finite subset of $\mathscr{A}$, not containing $\boldsymbol{X}$. Say

$$\mathscr{Y} = \{\boldsymbol{I} - S_\omega(Y_i^+): i \leqslant n\}.$$

For $i \leqslant n$, $X \neq Y_i$ and therefore, by (1), $X^+ \nsubseteq Y_i^+$.

Let $g(i)$ be an element of $X^+ - Y_i^+$, and let

$$U = \{g(i): i \leqslant n\}.$$

U is a finite subset of X^+ and so is not in $\boldsymbol{X}$. By construction, for $i \leqslant n$, $U \nsubseteq Y_i^+$, and so is in $\boldsymbol{I} - S_\omega(Y_i^+)$, and thus $U \in \bigcap \mathscr{Y}$. We have therefore shown that

$$\text{if } \boldsymbol{X} \in \mathscr{A}, \mathscr{Y} \in S_\omega(\mathscr{A}) \text{ and } \boldsymbol{X} \notin \mathscr{Y}, \text{ then } \bigcap \mathscr{Y} \nsubseteq \boldsymbol{X}. \tag{2}$$

Now let

$$\mathscr{I} = S_\omega(\boldsymbol{I}).$$

Again $\mathscr{I}$ is of cardinal α. Let

$$\mathfrak{B} = \{S_\omega(\boldsymbol{X}): \boldsymbol{X} \in \mathscr{A}\},$$

thus $\mathfrak{B} \subseteq \mathscr{P}(\mathscr{I})$ and $\mathfrak{B}$ is of cardinal 2^α.

Suppose

$$\mathfrak{X}=\{S_\omega(X_i): i\leqslant m\}$$

and

$$\mathfrak{Y}=\{S_\omega(Y_j): j\leqslant n\}$$

are disjoint finite subsets of $\mathfrak{B}$. Then for each $i\leqslant m$,

$$X_i\notin\{Y_j: j\leqslant n\}$$

and therefore, by (2), there is some element, say $h(i)$, in $\bigcap_{j\leqslant n}Y_j-X_i$. Let

$$\mathscr{Y}=\{h(i): i\leqslant m\}.$$

$\mathscr{Y}$ is a finite subset of each Y_j and therefore

$$\mathscr{Y}\in\bigcap_{j\leqslant n}S_\omega(Y_j)=\bigcap\mathfrak{Y}.$$

$\mathscr{Y}$ is not a subset of any X_i and therefore $\mathscr{Y}\notin\bigcup\mathfrak{X}$. Therefore we have shown that

$$\text{if } \mathfrak{X},\mathfrak{Y} \text{ are disjoint finite subsets of } \mathfrak{B},\ \bigcap\mathfrak{Y}\nsubseteq\bigcup\mathfrak{X}. \tag{3}$$

If $\mathfrak{Z}\subseteq\mathfrak{B}$, let

$$\mathfrak{Z}^*=(\mathfrak{B}-\mathfrak{Z})\cup\{\mathscr{I}-\mathscr{X}:\mathscr{X}\in\mathfrak{Z}\}.$$

If $\mathfrak{Z}_1,\mathfrak{Z}_2\in\mathfrak{B}$ and $\mathfrak{Z}_1\neq\mathfrak{Z}_2$, there is clearly some $\mathscr{X}\subseteq\mathscr{I}$ with $\mathscr{X}\in\mathfrak{Z}_1{}^*$ and $\mathscr{I}-\mathscr{X}\in\mathfrak{Z}_2{}^*$. Thus there are 2^{2^α} distinct sets $\mathfrak{Z}^*$, one corresponding to each subset of $\mathfrak{B}$.

Suppose $\mathfrak{Z}\subseteq\mathfrak{B}$ and let

$$\{\mathscr{X}_1,\ldots,\mathscr{X}_m,\mathscr{I}-\mathscr{Y}_1,\ldots,\mathscr{I}-\mathscr{Y}_n\}$$

be a finite subset of $\mathfrak{Z}^*$, with $\mathscr{X}_i\notin\mathfrak{Z}$ for $i\leqslant m$, and $\mathscr{Y}_j\in\mathfrak{Z}$ for $j\leqslant n$. By (3)

$$\mathscr{X}_1\cap\cdots\cap\mathscr{X}_m\nsubseteq\mathscr{Y}_1\cup\cdots\cup\mathscr{Y}_n,$$

whence

$$\mathscr{X}_1\cap\cdots\cap\mathscr{X}_m\cap(\mathscr{I}-\mathscr{Y}_1)\cap\cdots\cap(\mathscr{I}-\mathscr{Y}_n)\neq\emptyset.$$

This shows that $\mathfrak{Z}^*$ has the finite intersection property and therefore can be extended to an ultrafilter, say $U_\mathfrak{Z}$ on $\mathscr{I}$. By the remark above if $\mathfrak{Z}_1\neq\mathfrak{Z}_2$, $U_{\mathfrak{Z}_1}\neq U_{\mathfrak{Z}_2}$ and thus $\{U_\mathfrak{Z}:\mathfrak{Z}\subseteq\mathfrak{B}\}$ is a collection of 2^{2^α} ultrafilters on $\mathscr{I}$.

Let θ be any one–one map of $\mathscr{I}$ onto I. For each $\mathfrak{Z}\subseteq\mathfrak{B}$ let

$$V_\mathfrak{Z}=\{\theta[\mathfrak{X}]:\mathfrak{X}\in U_\mathfrak{Z}\}.$$

Then $\{V_\mathfrak{Z}:\mathfrak{Z}\subseteq\mathfrak{B}\}$ is a collection of 2^{2^α} ultrafilters on I. Since at most α of

these can be principal ultrafilters this shows that there are at least, and so exactly, 2^{2^α} non-principal ultrafilters on I.

(c). *Uniform ultrafilters.* An ultrafilter F on I is said to be *uniform* if all the sets in F are of the same cardinal as I.

LEMMA 1.6. *If I is infinite there is a uniform ultrafilter on I.*

PROOF. Let I be of cardinal α and let $\bar{S}_\alpha(I)$ be the set of all those subsets of I whose complements are of cardinal $<\alpha$. Thus $X\in\bar{S}_\alpha(I)$ iff $X\subseteq I$ and $\text{card}(I-X)<\alpha$. It is easily checked that $\bar{S}_\alpha(I)$ has the finite intersection property and so can be extended to an ultrafilter F on I. Clearly this ultrafilter contains no sets of cardinal $<\alpha$ and so is uniform.

Indeed any uniform ultrafilter on I must contain all the sets in $\bar{S}_\alpha(I)$, and conversely any ultrafilter containing $\bar{S}_\alpha(I)$ is uniform.

LEMMA 1.7. *Let I be an infinite set. The collections of non-principal ultrafilters on I and of uniform ultrafilters on I coincide iff I is countable.*

PROOF. All uniform ultrafilters are non-principal. If I is countable lemma 1.3 shows that any non-principal ultrafilter on I is uniform.

Suppose now $\text{card}(I)=\alpha>\aleph_0$. Let β be any infinite cardinal less than α and let J be any subset of I of cardinal β. Let

$$G=\{X\subseteq I: J-X \text{ is finite}\}.$$

G can easily be seen to have the finite intersection property and also to contain all the sets in $\bar{S}_\omega(I)$. Therefore G can be extended to a non-principal ultrafilter F on I. In particular $J\in G\subseteq F$, and so F is not uniform.

We will see in §3 that as far as ultraproducts are concerned we need only concern ourselves with uniform ultrafilters.

(d). *ω-incomplete ultrafilters.* An ultrafilter F on I is said to be *ω-complete* if whenever $\{X_n: n<\omega\}$ is a collection of elements of F then $\bigcap\{X_n: n<\omega\}$ is also in F. F is *ω-incomplete* if it is not ω-complete. This definition can be generalized to an arbitrary infinite cardinal α. F is said to be *α-complete* if whenever for $\xi<\alpha$, $X_\xi\in F$ then $\bigcap\{X_\xi:\xi<\alpha\}$ is also in F. F is *α-incomplete* if F is not *α-complete.*

LEMMA 1.8 *A principal ultrafilter is α-complete for all α.*

PROOF. If F is generated by $\{x\}$ then for any collection $\{X_\xi:\xi<\alpha\}$ of elements of F, $\{x\}\subseteq\bigcap\{X_\xi:\xi<\alpha\}$ and so $\bigcap\{X_\xi:\xi<\alpha\}\in F$.

The question naturally arises as to whether there exist non-principal ultrafilters which are ω-complete. The next easy lemma gives a complete answer in the case of ultrafilters on countable sets.

LEMMA 1.9. *If I is of cardinal α, there is no non-principal α-complete ultrafilter on I.*

PROOF. Let $\{x_\xi : \xi < \alpha\}$ be an enumeration of I and F a non-principal ultrafilter on I. For $\xi < \alpha$, let $X_\xi = I - \{x_\xi\}$. Since F is non-principal each set X_ξ is in F. But $\bigcap_{\xi<\alpha} X_\xi = \emptyset$ and so is not in F. It follows that F is not α-complete.

COROLLARY 1.10. *All non-principal ultrafilters on countable sets are ω-incomplete.*

The question as to whether ω-complete non-principal ultrafilters on uncountable sets exist is a famous problem in set theory. A cardinal α is said to be *measurable* if α is uncountable and there is a non-principal ultrafilter on α (or, equivalently, on some set of cardinal α) which is β-complete for each $\beta < \alpha$.

Readers familiar with measure theory will notice that if F is an ultrafilter on the set I, then the map

$$\mu : \mathscr{P}(I) \to \{0,1\},$$

defined by

$$\mu(X) = 1 \text{ if } X \in F$$

and

$$\mu(X) = 0 \text{ if } X \notin F$$

is a two valued measure on I which is non-trivial if F is non-principal and which is β-additive if F is β-complete. Thus α is a measurable cardinal iff α is uncountable and there is a non-trivial two valued measure on α which is β-additive for each $\beta < \alpha$. This explains the nomenclature.

We shall say that the cardinal α is *ω-measurable* if there is a non-principal ω-complete ultrafilter on α. In the literature the term 'measurable' is sometimes used for the notion we have called 'ω-measurable'. However the next result shows that as far as the existence of measurable cardinals is concerned this difference in nomenclature is unimportant.

THEOREM 1.11. *There exists a measurable cardinal iff there exists an ω-measurable cardinal.*

PROOF. Since all measurable cardinals are ω-measurable one way round is trivial. For the converse, suppose that there is an ω-measurable cardinal.

Let α be the least ω-measurable cardinal. By the previous lemma α must be uncountable.

Let F be a non-principal ω-complete ultrafilter on α. We show that for each $\beta<\alpha$, F is β-complete.

Suppose that F is not β-complete with $\beta<\alpha$. Then, as we show in the next lemma, there is a partition $\{X_\xi:\xi<\beta\}$ of α such that for each $\xi<\beta$, $X_\xi\notin F$.

Let $I=\{X_\xi:\xi<\beta\}$ and for $A\subseteq I$, let $\bigcup A=\bigcup\{X_\xi:X_\xi\in A\}$. We define the collection E of subsets of I by

$$E=\{A\subseteq I:\bigcup A\in F\}.$$

Clearly $I\in E$ and no finite subsets of I are in E. Suppose $A\in E$ and $A\subseteq B\subseteq I$. Then $\bigcup A\in F$ and $\bigcup A\subseteq\bigcup B$. Therefore $\bigcup B\in F$ and hence $B\in E$.

If $A\subseteq I$, $\bigcup A\cup\bigcup(I-A)=\alpha$. Thus either $\bigcup A\in F$ or $\bigcup(I-A)\in F$, but clearly not both. Hence either $A\in E$ or $I-A\in E$, but not both.

Finally suppose that for each $n\in\omega$, $A_n\in E$. Then for $n\in\omega$, $\bigcup A_n\in F$ and therefore, because F is ω-complete,

$$\bigcap_{n\in\omega}(\bigcup A_n)\in F.$$

But since the sets in I are pairwise disjoint,

$$\bigcap_{n\in\omega}(\bigcup A_n)=\bigcup(\bigcap_{n\in\omega}A_n),$$

and it follows that $\bigcap_{n\in\omega}A_n\in E$.

We have thus shown that E is an ω-complete non-principal ultrafilter on I. Since I is of cardinal $\beta<\alpha$ this contradicts our choice of α.

We can therefore conclude that F is β-complete for each $\beta<\alpha$ and so α is a measurable cardinal.

The existence of measurable cardinals remains an open problem. A most striking result is due to SCOTT [1961] who showed that the axiom of constructibility, that is $V=L$ in the sense of GÖDEL [1940], implies that measurable cardinals do not exist. We give a sketch of Scott's proof in ch. 14.

If I is a set, by a *partition* of I we mean a collection of pairwise disjoint non-empty sets whose union is I. We will make use of the following alternative characterizations of α-complete ultrafilters in the next section.

LEMMA 1.12. *Let F be an ultrafilter on I. F is α-complete iff whenever $\beta\leqslant\alpha$ and $\{X_\xi:\xi<\beta\}$ is a partition of I, then there is a unique $\xi_0<\beta$ such that $X_{\xi_0}\in F$.*

PROOF. Suppose F is α-complete, $\beta\leqslant\alpha$ and $\{X_\xi:\xi<\beta\}$ is a partition of I.

Since the X_ξ are pairwise disjoint at most one of them is in F. Suppose none of them is. Then for $\xi<\beta$, $I-X_\xi\in F$. Since F is α-complete

$$\bigcap_{\xi<\beta}(I-X_\xi)=I-\bigcup_{\xi<\beta}X_\xi=I-I=\emptyset\in F,$$

which contradicts the fact that F is proper.

The converse is proved similarly.

LEMMA 1.13. *Let F be an ultrafilter on I and let α be the least cardinal such that F is α-incomplete. Then there is a sequence $\langle X_\xi:\xi<\alpha\rangle$ of elements of F such that*

(1). $X_\xi\supseteq X_\eta$, *if* $\xi\leqslant\eta<\alpha$,

and

(2). $\bigcap_{\xi<\alpha}X_\xi=\emptyset$.

PROOF. Since F is α-incomplete there is a collection $\{Y_\xi:\xi<\alpha\}$ of elements of F such that $Y=\bigcap_{\xi<\alpha}Y_\xi\notin F$. Thus $I-Y\in F$. For $\xi<\alpha$, let

$$X_\xi=(I-Y)\cap\bigcap_{\zeta<\xi}Y_\zeta.$$

Since $\mathrm{card}(\xi)<\alpha$, F is $\mathrm{card}(\xi)$-complete and so $X_\xi\in F$. Clearly (1) holds and finally,

$$\bigcap_{\xi<\alpha}X_\xi=(I-Y)\cap\bigcap_{\xi<\alpha}Y_\xi=(I-Y)\cap Y=\emptyset.$$

This completes the proof.

(e). *Regular ultrafilters.* If I is infinite, $S_\omega(I)$, the set of all finite subsets of I, is of the same cardinal as I, and so there is a one–one map of I onto $S_\omega(I)$. We say that F is a *regular* ultrafilter on I if there is a one–one map f of I onto $S_\omega(I)$ such that for each $i\in I$,

$$\{j\in I:i\in f(j)\}\in F.$$

The importance of this concept will become apparent when we discuss the cardinality of ultraproducts.

We shall be interested in the relationship between regular ultrafilters and the other sorts of ultrafilters that we have described.

LEMMA 1.14. *Let I be an infinite set. Then there is a regular ultrafilter on I. Any regular ultrafilter on I is ω-incomplete and uniform.*

PROOF. Let f be a one–one map of I onto $S_\omega(I)$. For each $i\in I$ let

$$E_i=\{j\in I:i\in f(j)\}.$$

We show that the collection $E=\{E_i: i\in I\}$ has the finite intersection property.

Let $\{E_{i_1},\ldots,E_{i_n}\}$ be a finite subset of E. $\{i_1,\ldots,i_n\}$ is a finite subset of I and so is $f(i_0)$ for some $i_0\in I$. For $j\leqslant n$, $i_j\in f(i_0)$ and so

$$i_0\in E_{i_1}\cap\cdots\cap E_{i_n}.$$

Thus E has the finite intersection property and so can be extended to an ultrafilter F on I. By our construction F is a regular ultrafilter.

Now suppose that F is a regular ultrafilter on I. Then there is a one–one map f of I onto $S_\omega(I)$ such that for $i\in I$

$$\{j\in I: i\in f(j)\}\in F$$

Let $\{i_n: n\in\omega\}$ be a countably infinite subset of I. For $n\in\omega$, let

$$X_n=\{j\in I: i_n\in f(j)\}.$$

By hypothesis each X_n is in F, but obviously

$$\bigcap_{n\in\omega} X_n=\{j\in I: \text{for all } n\in\omega,\ i_n\in f(j)\}=\emptyset$$

and so is not in F. This shows that F is ω-incomplete.

Next we show that F is uniform. Suppose $X\subseteq I$ and $\operatorname{card}(X)<\operatorname{card}(I)$. Then, clearly,

$$\operatorname{card}\Big(\bigcup_{i\in X} f(i)\Big)<\operatorname{card}(I),$$

and therefore there is some $j_0\in I-\bigcup_{i\in X} f(i)$. If, for $i\in I$, $j_0\in f(i)$, $i\notin X$. It follows that

$$\{i\in I: j_0\in f(i)\}\subseteq I-X.$$

Since $\{i\in I: j_0\in f(i)\}\in F$, we can conclude that $I-X\in F$ and thus that $X\notin F$.

This shows that F is uniform.

It is not known, in general, whether all ω-incomplete uniform ultrafilters are regular. However once again we have a complete answer in the case of ultrafilters on countable sets, as in provided by the next two lemmas.

LEMMA 1.15. *Let $I=\{i_n: n\in\omega\}$ be a countable set and let F be a non-principal ultrafilter on I. There is a sequence $\langle X_n: n\in\omega\rangle$ of subsets of I such that*

(1). $X_0=I$,

(2). *for each n, $X_{n+1}\subseteq X_n$ and X_n-X_{n+1} is infinite,*

(3). *for each n, $X_n\in F$,*

(4). $\bigcap_{n\in\omega} X_n=\emptyset$.

PROOF. Let $X_0 = I$. Suppose $k < \omega$, and for $n \leqslant k$ we have defined X_n so as to satisfy (1)–(3) and

(5). for $m < n$, $i_m \notin X_n$.

Let A_k, B_k be infinite disjoint sets whose union is X_k. Since $X_k \in F$, either A_k or B_k is in F. If $A_k \in F$ then so also is $A_k - \{i_k\}$ and we let X_{k+1} be this set. If $A_k \notin F$, then $B_k \in F$ and so $B_k - \{i_k\} \in F$ and we let X_{k+1} be this set. Thus, in either case, we have found sets X_n for $n \leqslant k+1$ so that (1)–(3) and (5) are satisfied. Hence, by induction, for all n we can find a set X_n so that these conditions are satisfied. Since each X_n satisfies (5) it follows that $\bigcap_{n \in \omega} X_n = \emptyset$ and thus the proof is complete.

LEMMA 1.16. *If F is a non-principal ultrafilter on the countable set $I = \{i_n : n \in \omega\}$, then F is regular.*

PROOF. Let $\langle X_n : n \in \omega \rangle$ be the sequence of sets given by the previous lemma. For each $n \in \omega$, let Y_n be the collection of all those finite subsets of I which contain each i_k with $k < n$. Clearly,

(1). $Y_0 = S_\omega(I)$,

(2). for each n, $Y_{n+1} \subseteq Y_n$, and $Y_n - Y_{n+1}$ is infinite,

and

(3). $\bigcap_{n \in \omega} Y_n = \emptyset$.

For each n, let f_n be a one–one map of $X_n - X_{n+1}$ onto $Y_n - Y_{n+1}$ and let $f = \bigcup \{f_n : n \in \omega\}$. Because of the properties of the sets X_n and Y_n, f is a one–one map of I onto $S_\omega(I)$.

For each $n \in \omega$, if $i \in X_{n+1}$, $f(i) \in Y_{n+1}$ and therefore $i_n \in f(i)$. Hence

$$X_{n+1} \subseteq \{i \in I : i_n \in f(i)\},$$

but $X_{n+1} \in F$ and so $\{i \in I : i_n \in f(i)\} \in F$. This shows that F is regular.

LEMMA 1.17. *Let I be a countable set. Then the collections of non-principal, ω-incomplete, uniform and regular ultrafilters on I all coincide.*

PROOF. This is an immediate consequence of lemmas 1.7, 1.9 and 1.16.

We say that two ultrafilter pairs $\langle I, F \rangle$, $\langle J, G \rangle$ are *isomorphic*, if there is a one–one map π of I onto J such that for all $X \subseteq I$,

$$X \in F \quad \text{iff} \quad \pi[X] \in G.$$

If $\operatorname{card}(I) = \operatorname{card}(J)$, and F and G are both principal ultrafilters then clearly $\langle I, F \rangle$ is isomorphic to $\langle J, G \rangle$. Lemma 1.17 might tempt us to believe that this remains true if F and G are both non-principal ultrafilters. A simple cardinality argument shows that this would be a mistake.

LEMMA 1.18. *If I is infinite there are non-principal ultrafilters F, G on I such that $\langle I,F\rangle$ is not isomorphic to $\langle I,G\rangle$.*

PROOF. Let I be of cardinal α, and let F be a non-principal ultrafilter on I. We let Π be the set of all one–one maps of I onto I. Π is of cardinal 2^α. If $\pi\in\Pi$ let $\pi[F]=\{\pi[X]: X\in F\}$. Then $\{\langle I,\pi[F]\rangle : \pi\in\Pi\}$ is the collection of all ultrafilter pairs isomorphic to $\langle I,F\rangle$. This collection is of cardinal 2^α but by theorem 1.5 there are 2^{2^α} ultrafilter pairs $\langle I,G\rangle$ with G non-principal. Hence there is some (indeed 2^{2^α}) non-principal ultrafilter G on I such that $\langle I,F\rangle$ is not isomorphic to $\langle I,G\rangle$.

(f). *Good ultrafilters.* The motivation behind the definition of a good ultrafilter will not become apparent until ch. 11, however it is convenient to deal with them here. We first introduce some notation and nomenclature.

If α is a cardinal and X is a set,

$$S_\alpha(X)=\{Y\subseteq X: \operatorname{card}(Y)<\alpha\},$$
$$\bar{S}_\alpha(X)=\{Y\subseteq X: \operatorname{card}(X-Y)<\alpha\},$$

and

$$S^*(X)=\{Y\subseteq X: \operatorname{card}(Y)=\operatorname{card}(X)\}.$$

X is said to be *multiplicative* if X is closed under finite intersections. The *multiplicative closure* of X, denoted by $m(X)$, is the smallest multiplicative set containing X. Thus

$$m(X)=\{\bigcap Y: Y\in S_\omega(X)\}.$$

If f is a function Df is the domain of f and Rf is the range of f. If Y,Z are collections of sets, the function $f: Y\to Z$ is said to be

disjointed if $x,y\in Y$ and $x\neq y$ imply $f(x)\cap f(y)=\emptyset$,

monotonic if $x,y\in Y$ and $x\subseteq y$ imply $f(x)\subseteq f(y)$,

and

multiplicative if $x,y\in Y$ implies $f(x)\cap f(y)=f(x\cap y)$ whenever $x\cap y\in Y$.

If g,h are functions whose ranges are collections of sets we write $g\leqslant h$ if $Dg=Dh$ and for all x in Dg, $g(x)\subseteq h(x)$.

Let M be a multiplicative collection of sets. An ultrafilter F is said to be *M-good* if for each monotonic function $f: M\to F$ there is a multiplicative function $h: M\to F$ with $h\leqslant f$. If α is a cardinal, F is said to be *α-good* if F is $\bar{S}_\omega(\beta)$-good for all $\beta<\alpha$.

We let $Q(\alpha)$ be the collection of all α-good ultrafilters. Our aim is to show that $Q(\alpha^+)$ is large provided that $\alpha^+=2^\alpha$. To achieve this aim we proceed by a series of lemmas.

LEMMA 1.19. *Let β be an infinite cardinal and F an ultrafilter. F is β^+-good iff it is $\bar{S}_\omega(\beta)$-good.*

PROOF. By the definition, if F is β^+-good, F is $\bar{S}_\omega(\beta)$-good. For the converse suppose F is $\bar{S}_\omega(\beta)$-good and let γ be a cardinal $\leqslant\beta$. We need to show that F is $\bar{S}_\omega(\gamma)$-good.

Since β is infinite $\operatorname{card}(\bar{S}_\omega(\gamma))\leqslant\beta$. Let h be a map of β onto $\bar{S}_\omega(\gamma)$ and let $h^*:\bar{S}_\omega(\beta)\to\bar{S}_\omega(\gamma)$ be defined by

$$h^*(x)=\bigcap\{h(y):y\in\beta-x\}.$$

Then h^* is a monotonic and multiplicative function.

Now suppose $f:\bar{S}_\omega(\gamma)\to F$ is a monotonic function. Then $f^*:\bar{S}_\omega(\beta)\to F$ defined by

$$f^*(x)=f(h^*(x))$$

is a monotonic function, and therefore, by our assumption, there is a multiplicative function $g^*:\bar{S}_\omega(\beta)\to F$ with $g^*\leqslant f^*$. We define $g:\bar{S}_\omega(\gamma)\to F$ by

$$g(x)=\bigcup\{g^*(y):y\in\bar{S}_\omega(\beta)\text{ and }h^*(y)\subseteq x\}.$$

If $h^*(y)\subseteq x$,

$$g^*(y)\subseteq f^*(y)=f(h^*(y))\subseteq f(x),$$

since f is monotonic, and so $g(x)\subseteq f(x)$. Thus $g\leqslant f$. We conclude the proof by showing that g is multiplicative.

If $z\in g(x)\cap g(x')$, then for some $y,y'\in\bar{S}_\omega(\beta)$,

$$z\in g^*(y)\cap g^*(y'),\quad h^*(y)\subseteq x\quad\text{and}\quad h^*(y')\subseteq x'.$$

Because g^* is multiplicative $z\in g^*(y\cap y')$. Since $h^*(y)\subseteq x$ and $h^*(y')\subseteq x'$, and h^* is multiplicative, $h^*(y\cap y')\subseteq x\cap x'$. It follows that $z\in g(x\cap x')$.

Conversely if $z\in g(x\cap x')$, then for some $y\in\bar{S}_\omega(\beta)$,

$$z\in g^*(y)\quad\text{and}\quad h^*(y)\subseteq x\cap x'.$$

Since $h^*(y)\subseteq x$ and $h^*(y)\subseteq x'$, it follows that $z\in g(x)\cap g(x')$.

This completes the proof that $g(x)\cap g(x')=g(x\cap x')$, and thus of lemma 1.19.

From now on α will be a fixed but arbitrary infinite cardinal.

LEMMA 1.20. *Let $\{X_\xi:\xi<\alpha\}$ be a collection of α sets each of cardinal α. Then there is a collection $\{Y_\xi:\xi<\alpha\}$ of α pairwise disjoint sets each of cardinal α, with $Y_\xi\subseteq X_\xi$, for each $\xi<\alpha$.*

PROOF. We first sketch the idea involved in the proof. The case where we would seem to have the least chance of finding pairwise disjoint sets Y_ξ with the required properties would be where the X_ξ were all the same set X. But since α is infinite X can be expressed as the union of α disjoint sets each of cardinal α. These α sets form the desired sets Y_ξ. Although this is the case where we seemed least likely to succeed it is the easiest to handle as all elements are in the same set and so X can be decomposed any way we please. In the general case we have to be more careful as the precise proof below shows.

By the axiom of choice $\alpha \times \alpha = \alpha$, and so there is a one–one function h with domain α and range $\alpha \times \alpha$. Let F be the set of all one–one functions f such that $Df \in \alpha$ and whenever $\xi \in Df$ and $h(\xi) = \langle \zeta, \eta \rangle$ then $f(\xi) \in X_\zeta$. By the Hausdorff maximal principle F contains a maximal chain G with respect to set inclusion. Clearly $\bigcup G = g$ is a one–one function such that $Dg \in \alpha + 1$ and whenever $\xi \in Dg$ and $h(\xi) = \langle \zeta, \eta \rangle$ then $g(\xi) \in X_\zeta$.

Suppose $Dg \in \xi < \alpha$ and $h(\xi) = \langle \zeta, \eta \rangle$. Since card $(X_\zeta) = \alpha$ and card $(Rg) \leqslant$ $\leqslant$ card $(\xi) < \alpha$ there is some $x \in X_\zeta - Rg$. The function $g' = g \cup \{\langle \xi, x \rangle\}$ then belongs to F and $G \cup \{g'\}$ is a longer chain than G. This contradicts the choice of G and so $Dg = \alpha$.

Now for each $\zeta < \alpha$ let

$$Y_\zeta = \{g(\xi) : \xi < \alpha \text{ and for some } \eta < \alpha,\ h(\xi) = \langle \zeta, \eta \rangle\}.$$

Then it is easily seen that the collection $\{Y_\xi : \xi < \alpha\}$ has the required properties.

COROLLARY 1.21. *Let f be a function defined on α such that for $\xi < \alpha$, $f(\xi)$ is a set of cardinal α. Then there is a disjointed function g such that $g \leqslant f$ and for $\xi \leqslant \alpha$, $g(\xi)$ is a set of cardinal α.*

PROOF. This is just a restatement of lemma 1.20.

LEMMA 1.22. *Let D be a multiplicative subset of $S^*(\alpha)$ of cardinality $\leqslant \alpha$. Then there is a multiplicative subset E of $S^*(\alpha)$ of cardinal $\leqslant \alpha$ which contains D and is such that $\bigcap Y = \emptyset$ for some $Y \in S_{\omega^+}(E)$.*

PROOF. Let h be a map from α onto D. By corollary 1.21 there is a disjointed function $k \leqslant h$ such that $k(\xi)$ is of cardinal α for each $\xi < \alpha$. Clearly for each $\xi < \alpha$, we can find a sequence

$$k(\xi) = X_{\xi 0} \supseteq X_{\xi 1} \supseteq \cdots \supseteq X_{\xi n} \supseteq \cdots$$

of sets of cardinal α whose intersection is empty.

For each $n < \omega$, let $Y_n = \bigcup_{\xi < \alpha} X_{\xi n}$ and let $Y = \{Y_n : n \in \omega\}$. We will show that $\bigcap Y = \emptyset$.

Note first that

$$\bigcap Y = \bigcap_{n<\omega} \bigcup_{\xi<\alpha} X_{\xi n} = \bigcup_{f\in\alpha^\omega} \bigcap_{m<\omega} X_{f(m)m}.$$

For each $f\in\alpha^\omega$, if f is constant, then $\bigcap_{m<\omega} X_{f(m)m}=\emptyset$, by construction. While if f takes the two distinct values ξ,η,

$$\bigcap_{m<\omega} X_{f(m)m} \subseteq X_{\xi m}\cap X_{\eta m} \subseteq k(\xi)\cap k(\eta)=\emptyset,$$

by the choice of k. Thus $\bigcap Y=\emptyset$.

Now let $E=m(D\cup Y)$. E is a multiplicative set which is an extension of D and $Y\in S_{\omega+}(E)$. It remains to prove that $E\subseteq S^*(\alpha)$.

Suppose $z\in E$. Then, since D is multiplicative and because $Y_0\supseteq Y_1\supseteq\cdots$, for some $z'\in D$ and some $m<\omega$,

$$z=z'\cap Y_m.$$

Since h is an onto map, we also have that $z'=h(\eta)$ for some $\eta<\alpha$. Therefore

$$z=z'\cap Y_m = h(\eta)\cap \bigcup_{\xi<\alpha} X_{\xi m} \supseteq k(\eta)\cap \bigcup_{\xi<\alpha} X_{\xi m} = X_{\eta m}$$

and so is of cardinal α.

This completes the proof.

LEMMA 1.23. *Suppose $D\subseteq S^*(\alpha)$ is a multiplicative set of cardinal $\leqslant\alpha$, and $f:\bar{S}_\omega(\alpha)\to S^*(\alpha)$ is a monotonic function such that*

$$m(D\cup Rf)\subseteq S^*(\alpha).$$

Then there is a multiplicative function $g\leqslant f$ such that

$$m(D\cup Rg)\subseteq S^*(\alpha).$$

PROOF. Define $h:\bar{S}_\omega(\alpha)\times D\to S^*(\alpha)$ by

$$h(x,y)=f(x)\cap y.$$

Since

$$f(x)\cap y\in m(D\cup Rf),$$

$h(x,y)$ is always in $S^*(\alpha)$. Therefore, by corollary 1.21, there is a disjointed function $k\leqslant h$ which maps $\bar{S}_\omega(\alpha)\times D$ into $S^*(\alpha)$.

Define $g:\bar{S}_\omega(\alpha)\to S^*(\alpha)$ by

$$g(x)=\bigcup\{k(y,z): y\subseteq x,\ y\in\bar{S}_\omega(\alpha) \text{ and } z\in D\}.$$

Suppose $x\in\bar{S}_\omega(\alpha)$ and $\xi\in g(x)$; then there exist $y\subseteq x$, $y\in\bar{S}_\omega(\alpha)$, and $z\in D$ such that $\xi\in k(y,z)$. But

$$k(y,z)\subseteq h(y,z)=f(y)\cap z\subseteq f(x)\cap z\subseteq f(x),$$

and thus $\xi\in f(x)$. Therefore $g(x)\subseteq f(x)$ and so $g\leqslant f$.

Suppose $x,x'\in\bar{S}_\omega(\alpha)$ and $\xi\in g(x)\cap g(x')$. Then for some $y,y'\in\bar{S}_\omega(\alpha)$ with $y\subseteq x$, $y'\subseteq x'$, and some $z,z'\in D$

$$\xi\in k(y,z)\cap k(y',z').$$

k is disjointed and so this implies $y=y'$ and $z=z'$. Thus $y\subseteq x\cap x'$ and hence $\xi\in g(x\cap x')$. Conversely, if $\xi\in g(x\cap x')$, there is some $y\in\bar{S}_\omega(\alpha)$ such that $y\subseteq x,x'$ and some $z\in D$ with $\xi\in k(y,z)$. It follows at once that $\xi\in g(x)\cap g(x')$.

We have therefore shown that $g(x\cap x')=g(x)\cap g(x')$ and so g is multiplicative. To complete the proof we need only show that $m(D\cup Rg)\subseteq S^*(\alpha)$. Suppose $x\in m(D\cup Rg)$. Then since D and g are multiplicative, there is some $z\in D$ and some $y\in\bar{S}_\omega(\alpha)$ such that $x=g(y)\cap z$.

By the definition of $g(y)$, $k(y,z)\subseteq g(y)$ and since $k\leqslant h$,

$$k(y,z)\subseteq h(y,z)=f(y)\cap z\subseteq z,$$

so $k(y,z)\subseteq g(y)\cap z=x$. This shows that $x\in S^*(\alpha)$ and concludes the proof of the lemma.

We are now ready to prove the main result.

THEOREM 1.24. *Assume* $2^\alpha=\alpha^+$. *Then every multiplicative set* $D\subseteq S^*(\alpha)$ *of power* $\leqslant\alpha$ *can be extended to an* α^+*-good ultrafilter on* α.

PROOF. By lemma 1.22, D can be extended to a multiplicative subset of $S^*(\alpha)$ of cardinal $\leqslant\alpha$ which contains a countable subset Y with an empty intersection; let D_0 be this extension.

$\bar{S}_\omega(\alpha)$ is of cardinal α, and hence there are at most $(2^\alpha)^\alpha=2^\alpha=\alpha^+$ monotonic functions from $\bar{S}_\omega(\alpha)$ into $\mathscr{P}(\alpha)$. Let

$$\langle f_\xi:\xi<\alpha^+\rangle$$

be an enumeration of these functions.

By lemma 1.23, for each $\xi<\alpha^+$ and each multiplicative subset E of $S^*(\alpha)$ of cardinal $\leqslant\alpha$, such that $m(E\cup Rf_\xi)\subseteq S^*(\alpha)$, there is a multiplicative function $g_{\xi,E}$ such that

$$g_{\xi,E}\leqslant f_\xi$$

and

$$m(E\cup Rg_{\xi,E})\subseteq S^*(\alpha).$$

We define the sequence $\langle E_\xi : \xi < \alpha^+ \rangle$ by recursion as follows:

(1). $E_0 = D_0$,

(2). $E_\xi = \bigcup_{\zeta<\xi} E_\zeta$, if $\xi < \alpha^+$ is a limit ordinal,

(3). $E_{\xi+1} = E_\xi$, if $\xi < \alpha^+$ and $m(E_\xi \cup Rf_\xi) \not\subseteq S^*(\alpha)$,

(4). $E_{\xi+1} = m(E_\xi \cup Rg_{\xi,E_\xi})$, if $\xi < \alpha^+$ and $m(E_\xi \cup Rf_\xi) \subseteq S^*(\alpha)$.

Let $E_{\alpha^+} = \bigcup_{\xi<\alpha^+} E_\xi$. For each $\xi < \alpha^+$, $E_\xi \subseteq S^*(\alpha)$ and therefore $E_{\alpha^+} \subseteq S^*(\alpha)$. Also each E_ξ is multiplicative and hence so is E_{α^+}. It follows that $m(E_{\alpha^+} \cup \bar{S}_\alpha(\alpha))$ has the finite intersection property and so can be extended to a, necessarily uniform, ultrafilter F on α. Clearly $D \subseteq D_0 \subseteq E_{\alpha^+} \subseteq F$. Since $Y \subseteq D_0$, F is ω-incomplete.

We show that F is α^+-good. By lemma 1.19 it is sufficient to show that F is $\bar{S}_\omega(\alpha)$-good. Let $f : \bar{S}_\omega(\alpha) \to F$ be a monotonic function. Then for some $\xi < \alpha^+$, f is f_ξ. Since $Rf, E_\xi \subseteq F$ and F is a uniform filter

$$m(E_\xi \cup Rf) \subseteq S^*(\alpha).$$

g_{ξ,E_ξ} is a multiplicative function such that $g_{\xi,E_\xi} \leqslant f$ and by (4) $Rg_{\xi,E_\xi} \subseteq E_{\xi+1} \subseteq F$. Thus we can conclude that F is $\bar{S}_\omega(\alpha)$-good.

This completes the proof of the theorem.

COROLLARY 1.25. *If $2^\alpha = \alpha^+$, there is an α^+-good, ω-incomplete ultrafilter on α.*

PROOF. α is infinite and so can be expressed as the union

$$\alpha = \bigcup_{n \in \omega} Y_n$$

of $\aleph_0$ pairwise disjoint sets each of cardinal α.

For $m \in \omega$, let

$$X_m = \bigcup_{m \leqslant n \leqslant \omega} Y_n.$$

Then $E = \{X_m : m \in \omega\}$ is a multiplicative collection of subsets of $S^*(\alpha)$ of cardinal $\aleph_0$. Hence, by theorem 1.24, E can be extended to an α^+-good ultrafilter F on α. Since

$$\bigcap_{m \in \omega} X_m = \emptyset,$$

F is ω-incomplete.

§2. Set theoretic properties of ultraproducts

In this section we will be concerned with those properties of ultraproducts that only involve their algebraic or set theoretic structure, but not the language L.

Suppose $\langle I,F\rangle$ is an ultrafilter pair and that $G\subseteq I$. Then we let $F\restriction G=\{X\cap G: X\in F\}$.

LEMMA 2.1. *If $\langle I,F\rangle$ is an ultrafilter pair and $G\in F$, then $\langle G,F\restriction G\rangle$ is also an ultrafilter pair, and for all $X\subseteq I$,*

$$X\in F \text{ iff } X\cap G\in F\restriction G.$$

PROOF. The verification of this lemma is trivial and is left as an exercise for the reader.

THEOREM 2.2. *Let $\{\mathfrak{A}_i: i\in I\}$ be a family of relational structures, let $\langle I,F\rangle$ be an ultrafilter pair, and suppose $G\in F$, then*

$$\prod_{i\in I}\mathfrak{A}_i/F\simeq\prod_{i\in G}\mathfrak{A}_i/F\restriction G.$$

PROOF. We define the map

$$h:\prod_{i\in I}A_i/F\to\prod_{i\in G}A_i/F\restriction G$$

as follows:

If $f\in\prod_{i\in I}A_i$, then $f\restriction G\in\prod_{i\in G}A_i$ is the restriction of f to G. We let

$$h(f/F)=f\restriction G/F\restriction G.$$

Clearly h maps $\prod A_i/F$ onto $\prod A_i/F\restriction G$; we show that h is an isomorphism.

Let R be any n-ary relation of $\prod\mathfrak{A}_i/F$ and let S be the corresponding relation of $\prod\mathfrak{A}_i/F\restriction G$. Then by the definition of the ultraproduct, for any $f_1/F,\ldots,f_n/F\in\prod A_i/F$,

$$\begin{aligned}
\langle f_1/F,\ldots,f_n/F\rangle\in R \quad &\text{iff} \quad \{i\in I:\langle f_1(i),\ldots,f_n(i)\rangle\in R_i\}\in F\\
&\text{iff} \quad \{i\in I:\langle f_1(i),\ldots,f_n(i)\rangle\in R_i\}\cap G\in F\restriction G,\\
&\qquad \text{by lemma 2.1,}\\
&\text{iff} \quad \{i\in G:\langle f_1(i),\ldots,f_n(i)\rangle\in R_i\}\in F\restriction G\\
&\text{iff} \quad \langle f_1\restriction G/F\restriction G,\ldots,f_n\restriction G/F\restriction G\rangle\in S
\end{aligned}$$

that is

$$\text{iff} \quad \langle h(f_1/F),\ldots,h(f_n/F)\rangle\in S.$$

A similar argument shows that this equivalence also holds when R is the identity relation on $\prod A_i/F$ and S is the identity relation on $\prod A_i/F\restriction G$. Therefore h is an isomorphism between $\prod\mathfrak{A}_i/F$ and $\prod\mathfrak{A}_i/F\restriction G$, and the theorem is proved.

COROLLARY 2.3. *If F is a principal ultrafilter on I, then for some $k \in I$,*

$$\prod \mathfrak{A}_i/F \simeq \mathfrak{A}_k .$$

PROOF. Since F is principal for some $k \in I$, $\{k\} \in F$. By theorem 2.2.

$$\prod \mathfrak{A}_i/F \simeq \prod_{i \in \{k\}} \mathfrak{A}_i/F \restriction \{k\} \simeq \mathfrak{A}_k .$$

This corollary shows that we only get interesting results by considering non-principal ultrafilters.

COROLLARY 2.4. *Let F be a non-principal ultrafilter on I. There is a subset J of I and a uniform ultrafilter G on J such that for any family $\{\mathfrak{A}_i : i \in I\}$ of relational structures*

$$\prod_{i \in I} \mathfrak{A}_i/F \simeq \prod_{i \in J} \mathfrak{A}_i/G .$$

PROOF. Let α be the least cardinal of any set in F and let J be any set in F of cardinal α. By theorem 2.2 for any family $\{\mathfrak{A}_i : i \in I\}$ of relational structures

$$\prod_{i \in I} \mathfrak{A}_i/F \simeq \prod_{i \in J} \mathfrak{A}_i/G ,$$

where $G = F \restriction J$. For any $X \in F$, $X \cap J \in F$ and therefore, by the choice of α, $X \cap J$ is of cardinal α. Since all elements of G are of this form, G is a uniform ultrafilter on J.

Corollary 2.3 can be generalized in the following way.

LEMMA 2.5. *If $\mathfrak{A}$ is a relational structure of cardinal $\leqslant \alpha$ and F is an α-complete ultrafilter on I, then*

$$\mathfrak{A}^I/F \simeq \mathfrak{A} .$$

PROOF. It will be sufficient to prove that the canonical embedding $d : \mathfrak{A} \to \mathfrak{A}^I/F$ maps $\mathfrak{A}$ *onto* $\mathfrak{A}^I/F$.

Suppose $f/F \in A^I/F$. For each $a \in A$, let

$$X_a = \{i \in I : f(i) = a\} .$$

$\{X_a : a \in A\}$ is a partition of I into $\leqslant \alpha$ disjoint sets. Therefore, by lemma 1.12, for some $a_0 \in A$, $X_{a_0} \in F$. Hence $f/F = d(a_0)$. This shows that $A^I/F = d[A]$ and so completes the proof.

We conclude this section with two exercises which embody a result we need in the next chapter.

Suppose F is an ultrafilter on I and G is an ultrafilter on J. We let $F \otimes G$ be the collection of subsets of $I \times J$ given by

$$F \otimes G = \{X \subseteq I \times J : \{j \in J : \{i \in I : \langle i,j \rangle \in X\} \in F\} \in G\}$$

EXERCISE 2.6. Show that $F \otimes G$ is an ultrafilter on $I \times J$.

The definition of $F \otimes G$ may seem somewhat involved. The reader familiar with measure theory will see that it corresponds to the usual definition of the product measure. Intuitively, thinking of I and J as copies of the unit interval on the real line, $F \otimes G$ consists of all those subsets of the unit rectangle which intersect 'almost all' of the lines y=constant, in 'almost all' points.

EXERCISE 2.7. Show that

$$\prod_{j \in J} (\prod_{i \in I} \mathfrak{A}_{ij}/F)/G \simeq \prod_{\langle i,j \rangle \in I \times J} \mathfrak{A}_{ij}/F \otimes G,$$

where $\{\mathfrak{A}_{ij} : \langle i,j \rangle \in I \times J\}$ is any collection of relational structures indexed by $I \times J$.

(Hint: Define the map

$$\phi : \prod (\prod A_{ij}/F)/G \to \prod A_{ij}/F \otimes G$$

as follows. Suppose

$$f/G \in \prod (\prod A_{ij}/F)/G.$$

Then for each $j \in J$, $f(j) \in \prod A_{ij}/F$, say $f(j) = f_j/F$ with $f_j \in \prod_{i \in I} A_{ij}$. If $f^*(i,j) = f_j(i)$,

$$f^* \in \prod_{\langle i,j \rangle} A_{ij}.$$

Now put $\phi(f/G) = f^*/F \otimes G$ and show that ϕ is the required isomorphism.)

§ 3. The cardinality of ultraproducts

In this section we occupy ourselves with the question of how the cardinality of an ultraproduct depends on the cardinals of its factors, the cardinal of the index set and the nature of the ultrafilter. It will be seen that our classification of ultrafilters in § 1 is very relevant to the present discussion.

LEMMA 3.1. *Suppose that for each $i \in I$, $card(A_i) = card(B_i)$ and that F is an ultrafilter on I, then*

$$card(\prod A_i/F) = card(\prod B_i/F)$$

PROOF. Completely trivial.

Thus the cardinality of an ultraproduct depends only on the cardinals of the factors but not on exactly which sets they are. So we can assume they are cardinals. Thus we will be interested in trying to answer the question: if for $i \in I$, α_i is a cardinal and F is an ultrafilter on I, then what is the cardinal of $\prod \alpha_i/F$? For the most part, to avoid complications, we discuss only the case where the α_i are all equal, that is we will be interested in the cardinals of ultrapowers α^I/F.

Throughout this section I will be an arbitrary index set, of cardinal γ, for $i \in I$, α_i will be a cardinal, as will be α, and F will be an ultrafilter on I.

The first two lemmas are obvious.

LEMMA 3.2. *$min\{\alpha_i : i \in I\} \leqslant card(\prod \alpha_i/F) \leqslant \prod \alpha_i$.*

COROLLARY 3.3. *$\alpha \leqslant card(\alpha^I/F) \leqslant \alpha^\gamma$.*

COROLLARY 3.4. *If $\alpha = \alpha^\gamma$, $card(\alpha^I/F) = \alpha$.*

[In connection with this corollary we remark that, assuming the generalized continuum hypothesis, if $\gamma < \mathrm{cf}(\alpha)$ then $\alpha^\gamma = \alpha$. (See BACHMAN [1967] p. 173 for this and other results on cardinal arithmetic which are relevant here.)]

LEMMA 3.5. *For $i \in I$, let α_i and β_i be cardinals. If $\{i \in I : \beta_i \leqslant \alpha_i\} \in F$, then*

$$card(\prod \beta_i/F) \leqslant card(\prod \alpha_i/F).$$

PROOF. Let $X = \{i \in I : \beta_i \leqslant \alpha_i\}$. For $i \in X$, let ϕ_i be a one–one map of β_i into α_i and for $i \in I - X$ let a_i be an arbitrary fixed element of α_i. We define the map $\phi : \prod \beta_i/F \to \prod \alpha_i/F$ as follows:

Suppose $f \in \prod \beta_i$. We let $f^* \in \prod \alpha_i$ be the function such that

$$\text{for } i \in X, \qquad f^*(i) = \phi_i(f(i))$$

and

$$\text{for } i \in I - X, \quad f^*(i) = a_i.$$

Then we put $\phi(f/F) = f^*/F$. It is easily seen that ϕ is one–one and so the lemma is proved.

We now turn our attention to the cardinality of ultraproducts of finite sets.

LEMMA 3.6. *If α is a finite cardinal and F is any ultrafilter on I,*

$$card(\alpha^I/F) = \alpha.$$

PROOF. If α is finite, F is α-complete and so the result is an immediate consequence of lemma 2.5.

LEMMA 3.7. *Let $\{\alpha_i : i \in I\}$ be a collection of finite cardinals, and let F be an ultrafilter on I. If for some $n < \omega$, $\{i \in I : \alpha_i = n\} \in F$, then*

$$card(\prod \alpha_i/F) = n.$$

PROOF. For $n < \omega$, let $X_n = \{i \in I : \alpha_i = n\}$. Since the sets X_n are pairwise disjoint, at most one of them is in F. Suppose $X_n \in F$, then by theorem 2.2,

$$\prod_{i \in I} \alpha_i/F \simeq \prod_{i \in X_n} \alpha_i/F \restriction X_n.$$

and by lemma 3.6,

$$\text{card}\Big(\prod_{i \in X_n} \alpha_i/F \restriction X_n\Big) = n,$$

and so the lemma is proved.

COROLLARY 3.8. *If $\{\alpha_i : i \in I\}$ is a bounded set of finite cardinals and F is any ultrafilter on I,*

$$card(\prod \alpha_i/F) \text{ is finite}.$$

PROOF. Suppose that for all $i \in I$, $\alpha_i \leqslant n$. Then, with the notation of lemma 3.7, $I = X_1 \cup \ldots \cup X_n$. Hence one of the sets X_j, $1 \leqslant j \leqslant n$, is in F. The result now follows from lemma 3.7.

COROLLARY 3.9. *If $\{\alpha_i : i \in I\}$ is a collection of finite cardinals and F is an ω-complete ultrafilter on I, then*

$$card(\prod \alpha_i/F) \text{ is finite}.$$

PROOF. Since in this case one of the sets X_n must be in F.

The next lemma helps us to give an answer to our question when the hypothesis of lemma 3.7 does not apply.

LEMMA 3.10. *There exists a collection $A \subseteq \omega^\omega$ of integral valued functions of the natural numbers, of cardinal $2^{\aleph_0}$, such that*

(1). *if $f \in A$ and $n \in \omega$, $f(n) < 2^n$*

and

(2). *if $f, g \in A$ and $f \neq g$, then $\{n \in \omega : f(n) = g(n)\}$ is finite.*

PROOF. For each function $\phi \in 2^\omega$, we define the function $f_\phi \in \omega^\omega$ by

$$f_\phi(n) = \sum_{m < n} \phi(m) \cdot 2^m$$

It is clear that the family $A = \{f_\phi : \phi \in 2^\omega\}$ has the desired properties.

LEMMA 3.11. *Let $\{\alpha_i : i \in I\}$ be a collection of finite cardinals and let F be an ultrafilter on I such that none of the sets*

$$X_n = \{i \in I : \alpha_i = n\}$$

is in F. Then

$$2^{\aleph_0} \leqslant card(\prod \alpha_i / F).$$

PROOF. Since none of the X_n are in F, F must be ω-incomplete and therefore non-principal. Also it must be the case that there are arbitrarily large $n \in \omega$, with $X_n \neq \emptyset$. By taking suitable finite unions of those X_n which are not empty, we can therefore find a partition $\{J_n : n \in \omega\}$ of I, none of whose elements are in F, and such that if $i \in J_n$, $\alpha_i \geqslant 2^n$. For each $i \in I$, let $n(i)$ be the unique integer such that $i \in J_{n(i)}$.

For each $i \in I$, let $\langle m_{i,k} : k < 2^{n(i)} \rangle$ be a sequence of distinct elements of α_i, and let A be the family of functions in ω^ω given by the previous lemma.

For each $f \in A$, we define the element $h_f \in \prod \alpha_i$, by

$$h_f(i) = m_{i, f(n(i))}.$$

Since $f(n(i)) < 2^{n(i)}$, such an element always exists.

Suppose $f, g \in A$ and $h_f(i) = h_g(i)$. It follows that $f(n(i)) = g(n(i))$. Thus

$$\{i \in I : h_f(i) = h_g(i)\} \subseteq \bigcup \{J_n : f(n) = g(n)\}.$$

If $f \neq g$, there are only finitely many $n < \omega$, with $f(n) = g(n)$, and none of the J_n are in F. Therefore in this case

$$\bigcup \{J_n : f(n) = g(n)\} \notin F$$

and so

$$\{i \in I : h_f(i) = h_g(i)\} \notin F.$$

Hence $h_f/F \neq h_g/F$. Thus $\{h_f/F : f \in A\}$ is a subset of $\prod \alpha_i/F$ of cardinal $2^{\aleph_0}$. This completes the proof.

We are now able to give a complete answer to the question of the cardinality of ultraproducts of finite cardinals over a countable set.

THEOREM 3.12. *Let $\{\alpha_i : i \in I\}$ be a countable collection of finite cardinals, and let F be an ultrafilter on I. If for some $n \in \omega$, $\{i \in I : \alpha_i = n\} \in F$, then*

$$card(\prod \alpha_i / F) = n,$$

otherwise

$$card(\prod \alpha_i / F) = 2^{\aleph_0}.$$

PROOF. The first half of the theorem is lemma 3.7. If no $\{i \in I : \alpha_i = n\}$ is in F, by lemma 3.11, $2^{\aleph_0} \leqslant \operatorname{card}(\prod \alpha_i/F)$ and, by lemma 3.2,

$$\operatorname{card}(\prod \alpha_i/F) \leqslant \prod \alpha_i \leqslant \aleph_0{}^{\aleph_0} = 2^{\aleph_0},$$

and so the theorem is proved.

The more general question of determining for which infinite cardinals α there is a collection $\{\alpha_i : i \in I\}$ of finite cardinals and an ultrafilter F on I such that $\operatorname{card}(\prod \alpha_i/F) = \alpha$ is discussed by KEISLER [1967], under the assumption of the generalized continuum hypothesis.

This completes our discussion of ultraproducts of finite cardinals. The following remarks show that in our ensuing discussion of ultraproducts of arbitrary cardinals we can ignore the case when some of the factors are finite. For suppose $\{\alpha_i : i \in I\}$ is an arbitrary collection of cardinals and that F is an ultrafilter on I. Let $G = \{i \in I : \alpha_i \text{ is finite}\}$ and let $H = I - G$. Either G or H is in F, and so, by theorem 2.2, $\operatorname{card}(\prod_{i \in I} \alpha_i/F)$ is equal either to $\operatorname{card}(\prod_{i \in G} \alpha_i/F \restriction G)$ or to $\operatorname{card}(\prod_{i \in H} \alpha_i/F \restriction H)$. In the first case we have an ultraproduct of finite sets, which we have already dealt with. We now deal with the second case, where all the factors are infinite.

Our result for finite sets yields the following result for countable sets.

THEOREM 3.13. *If F is ω-complete*

$$\operatorname{card}(\omega^I/F) = \aleph_0$$

and if F is ω-incomplete

$$\operatorname{card}(\omega^I/F) \geqslant 2^{\aleph_0}.$$

In particular, if $\operatorname{card}(I) = \aleph_0$ and F is ω-incomplete,

$$\operatorname{card}(\omega^I/F) = 2^{\aleph_0}.$$

PROOF. When F is ω-complete, it follows at once from lemma 2.5 that $\operatorname{card}(\omega^I/F) = \aleph_0$.

Suppose now that F is ω-incomplete. Then by lemma 1.12 there is a partition $\{X_n : n \in \omega\}$ of I none of whose elements are in F. For $i \in X_n$, let $\alpha_i = n$. By lemma 3.11, $2^{\aleph_0} \leqslant \operatorname{card}(\prod \alpha_i/F)$ and by lemma 3.5,

$$\operatorname{card}(\prod \alpha_i/F) \leqslant \operatorname{card}(\omega^I/F).$$

This proves the result.

In particular, when $\operatorname{card}(I) = \aleph_0$, we have

$$2^{\aleph_0} \leqslant \operatorname{card}(\omega^I/F) \leqslant \aleph_0{}^{\aleph_0} = 2^{\aleph_0}.$$

COROLLARY 3.14. *If the cardinals α_i are all infinite and F is ω-incomplete*

$$2^{\aleph_0} \leqslant card\left(\prod \alpha_i/F\right).$$

It follows from this that, except in trivial cases, an ultraproduct is always uncountable.

The next result, due to Keisler, gives some more information about the cardinality of ultraproducts over ω-incomplete ultrafilters. We need first the following set theoretic lemma.

LEMMA 3.15. *If F is an ω-incomplete ultrafilter on I, there is a map*

$$f: I \to S_\omega(\omega)$$

such that for each $n \in \omega$,

$$\{i \in I : n \in f(i)\} \in F.$$

PROOF. Let $\langle X_n : n \in \omega \rangle$ be the descending sequence of elements of F with an empty intersection given by lemma 1.13; clearly we can assume $X_0 = I$. For each $i \in I$, let

$$f(i) = \{n \in \omega : i \in X_n\}.$$

Since $\bigcap X_n = \emptyset$, $f(i)$ is a finite subset of ω, that is, $f(i) \in S_\omega(\omega)$.

If $n \in \omega$, and $i \in X_n$, then $n \in f(i)$. Thus, for $n \in \omega$,

$$X_n \subseteq \{i \in I : n \in f(i)\}.$$

Since $X_n \in F$ we can conclude that

$$\{i \in I : n \in f(i)\} \in F.$$

THEOREM 3.16. *If F is an ω-incomplete ultrafilter, and each α_i is infinite,*

$$[card(\prod \alpha_i/F)]^\omega = card(\prod \alpha_i/F).$$

PROOF. Let $f: I \to S_\omega(\omega)$ be the function given by the previous lemma. Since $f(i)$ is finite, there is, for each $i \in I$, a one–one map g_i of ${\alpha_i}^{f(i)}$ onto α_i. We are going to define a map

$$\theta: (\prod \alpha_i/F)^\omega \to \prod \alpha_i/F.$$

Suppose

$$h = \langle h_n/F : n \in \omega \rangle \in (\prod \alpha_i/F)^\omega.$$

Then for each $n < \omega$, $h_n \in \prod \alpha_i$, and so for $i \in I$, $h_n(i) \in \alpha_i$. For $i \in I$, we define $h_i^* \in {\alpha_i}^{f(i)}$ by putting

$$h_i^*(n) = h_n(i) \quad \text{for each } n \in f(i).$$

Now we define $g \in \prod \alpha_i$ by putting

$$g(i) = g_i(h_i^*) \quad \text{for each } i \in I$$

and finally we put

$$\theta(h) = g/F .$$

Suppose h,h' are distinct elements of $(\prod \alpha_i/F)^\omega$,

$$\theta(h) = g/F \quad \text{and} \quad \theta(h') = g'/F .$$

Since $h \neq h'$, for some $n \in \omega$, $h_n/F \neq h'_n/F$ and so

$$X = \{i \in I : h_n(i) \neq h'_n(i)\} \in F .$$

Let $Y = \{i \in I : n \in f(i)\}$; by hypothesis $Y \in F$ and therefore $X \cap Y \in F$.

Suppose $i \in X \cap Y$. Then $h_n(i) \neq h'_n(i)$ and $n \in f(i)$. Therefore $h_i^*(n) \neq h_i'^*(n)$, so $h_i^* \neq h_i'^*$. Since g_i is one–one it follows that $g(i) \neq g'(i)$. Thus

$$X \cap Y \subseteq \{i \in I : g(i) \neq g'(i)\} = Z , \text{ say.}$$

Hence $Z \in F$ and we can conclude that $g/F \neq g'/F$. θ is therefore a one–one function and so

$$[\text{card}(\prod \alpha_i/F)]^\omega \leqslant \text{card}(\prod \alpha_i/F) .$$

Since the converse inequality is obvious, the theorem is proved.

COROLLARY 3.17. *If α is an infinite cardinal and F is an ω-incomplete ultrafilter on I,*

$$[card(\alpha^I/F)]^\omega = card(\alpha^I/F) .$$

Note that since $\aleph_0^{\aleph_0} \neq \aleph_0$ this gives us another proof that ultraproducts are rarely countable.

The next lemma enables us to obtain some information about ultraproducts over uniform ultrafilters.

LEMMA 3.18. *If γ is infinite and $\alpha \geqslant \gamma$ there is a subset X of α^I of cardinal $> \gamma$ such that*

$$\text{if } f,g \in X \text{ and } f \neq g \text{ then } card(\{i \in I : f(i) = g(i)\}) < \gamma \tag{1}$$

PROOF. Let $\{x_\eta : \eta < \gamma\}$ be an enumeration of I. Let Y be the collection of all subsets of α^I which satisfy (1). Clearly chains in Y with respect to set inclusion have upper bounds and so, by Zorn's lemma, Y contains a maximal element X. We shall prove that $\text{card}(X) > \gamma$.

Suppose not and that X can be enumerated as $\langle f_\xi : \xi < \gamma \rangle$. We define the function $g \in \alpha^I$ as follows.

First we put $g(x_0)=0$. Now suppose $\zeta<\gamma$ and that $g(x_\eta)$ has been defined for all $\eta<\zeta$. The set $\{f_\xi(x_\zeta):\xi<\zeta\}$ is of cardinal $<\gamma$. Therefore there is some element of α not in this set. Let λ_ζ be the least such element of α. Then we put $g(x_\zeta)=\lambda_\zeta$.

It follows from this definition that $g(x_\zeta)\neq f_\xi(x_\zeta)$ whenever $\xi<\zeta<\gamma$.

Now let f be any element of X, say $f=f_\mu$. By the remark above the set $\{i\in I: f_\mu(i)=g(i)\}\subseteq\{x_\xi:\xi\leqslant\mu\}$ and so is of cardinal $<\gamma$. Thus $X\cup\{g\}$ satisfies (1). This contradicts our choice of X as a maximal element of Y and this contradiction shows that $\operatorname{card}(X)>\gamma$ thus proving the lemma.

THEOREM 3.19. *If $\alpha\geqslant\gamma\geqslant\aleph_0$ and F is a uniform ultrafilter on I, then*

$$\operatorname{card}(\alpha^I/F)>\gamma$$

PROOF. Let X be the family given by the previous lemma and let

$$A=\{f/F:f\in X\}.$$

If $f,g\in X$ and $f\neq g$, $\operatorname{card}(\{i\in I:f(i)=g(i)\})<\gamma$, and therefore, because F is a uniform ultrafilter,

$$\{i\in I:f(i)=g(i)\}\notin F$$

and so $f/F\neq g/F$. Since $\operatorname{card}(X)>\gamma$, the theorem is proved.

EXAMPLE 3.20. Let I be of cardinal $\aleph_0$, let J be of cardinal $2^{\aleph_0}$ and let F be a uniform ultrafilter on J. By theorem 3.19,

$$\operatorname{card}((2^I)^J/F)>2^{\aleph_0}.$$

But $\operatorname{card}(2^J/F)=2$ and so

$$\operatorname{card}((2^J/F)^I)=2^{\aleph_0},$$

and therefore,

$$(2^I)^J/F\neq(2^J/F)^I.$$

This shows that the operation of forming direct powers does not commute with that of forming ultraproducts.

We conclude this section with a result which gives us complete information about the cardinality of ultrapowers over regular ultrafilters.

THEOREM 3.21. *If α is infinite and F is a regular ultrafilter on I,*

$$\operatorname{card}(\alpha^I/F)=\alpha^\gamma.$$

PROOF. By corollary 3.3 it will be sufficient to prove that $\alpha^\gamma\leqslant\operatorname{card}(\alpha^I/F)$.

Let A be the set of all finite sequences of elements of α. Since α is infinite $\text{card}(A)=\alpha$, therefore it will suffice to show that $\alpha^\gamma \leqslant \text{card}(A^I/F)$ and we shall do this by constructing a one–one map of α^I into A^I/F.

Since F is regular there is a one–one map ϕ of I into $S_\omega(I)$ such that for each $j\in I$,

$$J_j=\{i\in I: j\in\phi(i)\}\in F.$$

Let $<$ be some fixed linear ordering of I. Suppose $f\in\alpha^I$. We construct the function $f^*\in A^I$ as follows. For $i\in I$, $\phi(i)$ is a finite subset of I, say $\{i_0,\ldots,i_k\}$ with $i_0<i_1<\cdots<i_k$. Then we put

$$f^*(i)=\langle f(i_0),\ldots,f(i_k)\rangle,$$

and hence $f^*(i)\in A$. Now we let

$$h(f)=f^*/F.$$

It remains only to show that h is one–one.

Suppose $f,g\in\alpha^I$ and $f\neq g$. Then for some $j\in I$, $f(j)\neq g(j)$. Now let i be an element of I such that $j\in\phi(i)$, say

$$\phi(i)=\{i_0,\ldots,j,\ldots,i_k\},$$

with $i_0<\cdots<j<\cdots<i_k$. Then

$$f^*(i)=\langle f(i_0),\ldots,f(j),\ldots,f(i_k)\rangle,$$

and

$$g^*(i)=\langle g(i_0),\ldots,g(j),\ldots,g(i_k)\rangle,$$

and therefore $f^*(i)\neq g^*(i)$. Thus we have shown that

$$J_j=\{i\in I: j\in\phi(i)\}\subseteq\{i\in I: f^*(i)\neq g^*(i)\}.$$

But $J_j\in F$, and therefore, since F is a filter,

$$\{i\in I: f^*(i)\neq g^*(i)\}\in F.$$

It follows that $f^*/F\neq g^*/F$, that is, that $h(f)\neq h(g)$. Hence h is one–one and our proof is completed.

§4. Well-ordered ultraproducts.

For each $i\in I$, let $\mathfrak{A}_i=\langle A_i,R_i\rangle$ be an ordered set. That is, let R_i be a total ordering of the set A_i. We say that R_i is a *well-ordering* if there is no sequence $\langle a_n: n<\omega\rangle$ of distinct elements of A_i such that for each $n<\omega$, $\langle a_{n+1},a_n\rangle\in R_i$.

Or, writing $a_n >_i a_{n+1}$ for $\langle a_{n+1}, a_n \rangle \in R_i$, if there is no infinite descending sequence

$$a_0 >_i a_1 >_i \cdots >_i a_n >_i a_{n+1} >_i \cdots$$

of elements of A_i. It is easily seen that R_i is a well-ordering iff each non-empty subset of A_i has a least element, with respect to the ordering R_i. Thus our definition coincides with the usual one.

From now on we assume that each R_i is a well-ordering and we answer the question: in what circumstances is the ultraproduct $\prod \mathfrak{A}_i/F$ well-ordered?

THEOREM 4.1. *If F is an ω-complete ultrafilter, then $\prod \mathfrak{A}_i/F$ is well-ordered.*

PROOF. Suppose that $\prod \mathfrak{A}_i/F$ is not well-ordered and that

$$f_0/F >_F f_1/F >_F \cdots >_F f_n/F >_F f_{n+1}/F >_F \cdots$$

is an infinite descending sequence in $\prod \mathfrak{A}_i/F$. Then for each $n \in \omega$ we have

$$X_n = \{i \in I : f_n(i) >_i f_{n+1}(i)\} \in F.$$

Since F is ω-complete

$$X = \bigcap_{n<\omega} X_n$$

is in F and so is not empty. But if i is any element of X

$$f_0(i) >_i f_1(i) >_i \cdots >_i f_n(i) >_i f_{n+1}(i) >_i \cdots$$

is an infinite descending sequence of elements of A_i, contradicting the fact that $\mathfrak{A}_i$ is well-ordered by R_i. This contradiction proves the theorem.

THEOREM 4.2. *Suppose each $\mathfrak{A}_i$ is finite, then $\prod \mathfrak{A}_i/F$ is well-ordered iff it is finite.*

PROOF. Trivially, if $\prod \mathfrak{A}_i/F$ is finite, it is well-ordered. If not, then an argument similar to that used in proving theorem 5.3.9 shows that $\prod \mathfrak{A}_i/F$ is not well-ordered.

THEOREM 4.3. *If each $\mathfrak{A}_i$ is infinite and F is not ω-complete then $\prod \mathfrak{A}_i/F$ is not well-ordered.*

PROOF. This theorem can be deduced from the previous one in a manner similar to that used to deduce theorem 3.13 from lemma 3.11. However we prefer to give a more direct proof.

Each $\mathfrak{A}_i$, being an infinite well-ordered set, has an initial sequence of order type ω. We can assume that in each case this initial sequence is

$$0 <_i 1 <_i 2 <_i \cdots <_i n <_i \cdots$$

Let $\langle X_n : n<\omega\rangle$ be the sequence of elements of F given by lemma 1.13; clearly we can assume $X_0 = I$. Thus, for each $n<\omega$, $X_n \supseteq X_{n+1}$ and $\bigcap X_n = \emptyset$. For each $i \in I$, let

$$\tau(i) = \max\{n : i \in X_n\}.$$

We define the sequence $\langle f_n : n<\omega\rangle$ of elements of $\prod A_i$ as follows:

For $i \in I$,

$$f_0(i) = \tau(i),$$

and

$$f_{n+1}(i) = f_n(i) \dot{-} 1,$$

where $k \dot{-} 1 = k-1$ if $k>0$, and $0 \dot{-} 1 = 0$. Thus

$$f_{n+1}(i) <_i f_n(i) \quad \text{iff} \quad n+1 \leqslant \tau(i) \quad \text{iff} \quad i \in X_{n+1},$$

and therefore

$$\{i \in I : f_{n+1}(i) <_i f_n(i)\} = X_{n+1} \in F,$$

and so

$$f_0/F >_F f_1/F >_F \cdots >_F f_n/F >_F f_{n+1}/F >_F \cdots$$

is an infinite descending sequence of elements of $\prod \mathfrak{A}_i/F$ which is therefore not well-ordered.

In particular, if for $n<\omega$, $\mathfrak{A}_n$ is the set of natural numbers with their usual ordering, and F is a non-principal ultrafilter on ω, notice that

$$\begin{aligned} \langle 0,1,2,3,4,5,6,\ldots\rangle/F &>_F \langle 0,0,1,2,3,4,5,\ldots\rangle/F \\ &>_F \langle 0,0,0,1,2,3,4,\ldots\rangle/F \\ &>_F \langle 0,0,0,0,1,2,3,\ldots\rangle/F \\ &>_F \cdots. \end{aligned}$$

is an infinite descending sequence in $\prod \mathfrak{A}_n/F$.

In view of theorem 2.2, the theorems of this section give a complete answer to our question about well-ordered ultraproducts. It can be seen that, unless ω-complete non-principal ultrafilters exist, ultraproducts are only well-ordered in trivial cases.

A generalization of the notion of a *well-ordering* is the following. A relation R on a set X is said to be *well-founded* if there is no sequence $\langle x_n : n<\omega\rangle$ of distinct elements of X such that for each $n<\omega$,

$$\langle x_{n+1}, x_n\rangle \in R.$$

Thus a well-ordering is a total ordering that is well-founded. It is immediate that all the results of this section extend to well-founded relations.

We shall make use of the results proved here in subsequent chapters.

§5. The Rabin–Keisler theorem

This is a convenient place to consider the following problem. Let α be an infinite cardinal. Does every relational structure of cardinal α have a proper elementary extension of cardinal α? The first result relevant to this question is a straightforward application of ultrapowers.

THEOREM 5.1. *If $\alpha=\alpha^{\aleph_0}$, then every relational structure of cardinal α has a proper elementary extension of cardinal α.*

PROOF. Suppose $\alpha=\alpha^{\aleph_0}$, and let $\mathfrak{A}$ be a relational structure of cardinal α. Let F be a non-principal ultrafilter on ω. By lemma 5.2.3 the ultrapower $\mathfrak{A}^\omega/F$ is, up to isomorphism, an elementary extension of $\mathfrak{A}$. Let $a^*=\langle a_0,a_1,\ldots,a_n,\ldots\rangle$ be a sequence of distinct elements of A. Since F is non-principal, for each $a_+\in A$,

$$d(a_+)\neq a^*/F,$$

where d is the canonical embedding of $\mathfrak{A}$ into $\mathfrak{A}^\omega/F$. It follows that $\mathfrak{A}^\omega/F$ is a proper elementary extension of $d[\mathfrak{A}]$. Now

$$\alpha\leqslant\operatorname{card}(\mathfrak{A}^\omega/F)\leqslant\alpha^{\aleph_0}=\alpha$$

and $d[\mathfrak{A}]\simeq\mathfrak{A}$. This proves the theorem.

We now show that in case α is not ω-measurable, the condition $\alpha=\alpha^{\aleph_0}$ is also necessary for the truth of the result above. We need first the following result from set theory which can be regarded as a generalization of lemma 3.10.

LEMMA 5.2. *There is a collection of $\alpha^{\aleph_0}$ countably infinite subsets of α, whose pairwise intersections are finite.*

PROOF. Since α is infinite there is a one–one function ϕ mapping the set of all finite sequences of elements of α onto α. If $f\in\alpha^\omega$, let

$$R_f=\{\phi(\langle f(0),\ldots,f(k)\rangle):k\in\omega\}.$$

R_f is a countably infinite subset of α. Suppose $f,g\in\alpha^\omega$ and $f\neq g$. Then for

some $n<\omega$, $f(n)\neq g(n)$. Therefore if $k\geqslant n$

$$\langle f(0),\ldots,f(n),\ldots,f(k)\rangle\neq\langle g(0),\ldots,g(n),\ldots,g(k)\rangle$$

and so

$$\phi(\langle f(0),\ldots,f(k)\rangle)\neq\phi(\langle g(0),\ldots,g(k)\rangle).$$

It follows that $R_f\cap R_g$ is finite, and thus that the collection $\{R_f: f\in\alpha^\omega\}$ has the desired properties.

We need also the following result.

LEMMA 5.3. *Suppose* $\mathfrak{A}=\langle A,\{R_\xi:\xi<\beta\}\rangle$ *where* $\{R_\xi:\xi<\beta\}$ *is an enumeration of all unary relations on* A *(i.e. all subsets of* A*) and that* $\mathfrak{A}'=\langle A',\{R'_\xi:\xi<\beta\}\rangle$ *is a proper elementary extension of* $\mathfrak{A}$. *If* $x\in A'-A$ *and*

$$F=\{R_\xi: x\in R'_\xi\}$$

then F *is a non-principal ultrafilter on* A.

PROOF. The appropriate language for the structures $\mathfrak{A},\mathfrak{A}'$ has a collection $\{P_\xi:\xi<\beta\}$ of unary predicate letters, corresponding to the unary relations $\{R_\xi:\xi<\beta\}$ and $\{R'_\xi:\xi<\beta\}$.

Suppose first that $R_\xi\in F$ and $R_\xi\subseteq R_\zeta$. Then

$$\mathfrak{A}\models(\forall v)[P_\xi(v)\rightarrow P_\zeta(v)],$$

but $\mathfrak{A}\prec\mathfrak{A}'$ and so

$$\mathfrak{A}'\models(\forall v)[P_\xi(v)\rightarrow P_\zeta(v)].$$

Hence $R'_\xi\subseteq R'_\zeta$. Since $R_\xi\in F$, $x\in R'_\xi$. Therefore $x\in R'_\zeta$ and so $R_\zeta\in F$.

In a similar way we can show that if R_ξ and R_ζ are in F then so also is $R_\xi\cap R_\zeta$, and that for each $\xi<\beta$ either R_ξ or $A-R_\xi$ is in F. This shows that F is an ultrafilter on A.

If F were a principal ultrafilter it would contain a one element set, say R_ξ. Then $\mathfrak{A}\models(\exists!^1 v)P_\xi(v)$ and so $\mathfrak{A}'\models(\exists!^1 v)P_\xi(v)$, whence R'_ξ is also a one element set. Since $R_\xi\in F$, $x\in R'_\xi$, and because $\mathfrak{A}\subseteq\mathfrak{A}'$, $R_\xi\subseteq R'_\xi$. It follows that $R_\xi=\{x\}=R'_\xi$. Since this contradicts the fact that $x\notin A$ we can conclude that F is non-principal.

THEOREM 5.4. *If* α *is not* ω*-measurable and all relational structures of cardinal* α *have proper elementary extensions of cardinal* α, *then* $\alpha=\alpha^{\aleph_0}$.

PROOF. Let A be a set of cardinal α and let

$$\mathfrak{A}=\langle A,\{R_\xi:\xi<\beta\}\cup\{S_\xi:\xi<\beta\}\rangle$$

where $\{R_\xi : \xi < \beta\}$ is an enumeration of all unary relations on A and $\{S_\xi : \xi < \beta\}$ is an enumeration of all binary relations on A.

Let $\{R_f : f \in \alpha^\omega\}$ be the family of subsets of A whose existence is given by lemma 5.2. For each $f \in \alpha^\omega$ there is some $\xi_f < \beta$ such that $R_f = R_{\xi_f}$. If $f \in \alpha^\omega$ we let

$$\langle r_{f,n} : n < \omega \rangle$$

be an enumeration of the countably infinite set R_f.

By hypothesis $\mathfrak{A}$ has a proper elementary extension of cardinal α, say

$$\mathfrak{A}' = \langle A', \{R'_\xi : \xi < \beta\} \cup \{S'_\xi : \xi < \beta\} \rangle$$

Let x be an element of A' not in A and let

$$F = \{R_\xi : x \in R'_\xi\}.$$

By lemma 5.3, F is a non-principal ultrafilter on A. Therefore since α is not ω-measurable, F is ω-incomplete. Hence, by lemma 1.13, there is a descending sequence

$$A = F_0 \supseteq F_1 \supseteq \cdots \supseteq F_n \supseteq \cdots$$

of elements of F whose intersection is empty. For each $n < \omega$, there is some $\xi_n < \beta$ such that $F_n = R_{\xi_n}$. We let $F'_n = R'_{\xi_n}$, and for each $f \in \alpha^\omega$, we define the function $\tau_f : A \to R_f$ by

$$\tau_f(a) = r_{f,n} \quad \text{iff} \quad a \in F_n - F_{n+1}.$$

The relation

$$T_f = \{\langle a, b \rangle : \tau_f(a) = b\}$$

is one of the relations S_ξ, say S_{ξ_f}. Therefore using the fact that $\mathfrak{A} \prec \mathfrak{A}'$ in a way similar to that adopted in proving lemma 5.3, it follows that τ'_f defined by

$$\tau'_f(a') = b' \quad \text{iff} \quad \langle a', b' \rangle \in S'_{\xi_f},$$

is a function from A' into R'_f and that

$$\tau'_f(a') = r_{f,n} \quad \text{iff} \quad a' \in F'_n - F'_{n+1}.$$

Since for all $n < \omega$, $x \in F'_n$, $\tau'_f(x) \notin R_{\xi_f}$ and so

$$\tau'_f(x) \in R'_{\xi_f} - R_{\xi_f}.$$

Similarly, since when $f \neq g$, $R_{\xi_f} \cap R_{\xi_g}$ is finite, it can be seen that

$$R_{\xi_f} \cap R_{\xi_g} = R'_{\xi_f} \cap R'_{\xi_g}.$$

It follows from these last two remarks that when $f \neq g$,

$$\tau'_f(x) \neq \tau'_g(x),$$

and hence A' contains at least $\alpha^{\aleph_0}$ elements. Since $\mathfrak{A}'$ is of cardinal α, we can conclude that $\alpha = \alpha^{\aleph_0}$ and our proof is complete.

It should be noted that the relational structure $\mathfrak{A}$ used in this proof has more than α (in fact 2^α) relations. It is a simple consequence of the completeness theorem that all relational structures of cardinal α, having at most α relations, have proper elementary extensions of cardinal α, irrespective of whether or not $\alpha = \alpha^{\aleph_0}$.

EXERCISE 5.5. Verify this last remark.

Let $\mathfrak{A}$ be the relational structure defined in the proof of theorem 5.4. Notice that it follows from the proof of this theorem that $\mathfrak{A}$ does not have an elementary extension of cardinal γ, if $\alpha < \gamma < \alpha^{\aleph_0}$. This example explains why we had to mention the cardinal of the language L in our statement of theorem 4.3.3. It has recently been shown by Keisler that a relational structure of cardinal α has an elementary extension of all cardinals $\geqslant \alpha^{\aleph_0}$ irrespective of the cardinal of the language. For a proof of this see MORLEY [1968].

Putting together theorems 5.1 and 5.4 we at once obtain

THEOREM 5.6. THE RABIN–KEISLER THEOREM
Let α be an infinite cardinal which is not ω-measurable. All relational structures of cardinal α have proper elementary extensions of cardinal α iff $\alpha = \alpha^{\aleph_0}$.

§ 6. Historical and bibliographical remarks

Theorem 1.5 is due to TARSKI [1939]. Uniform ultrafilters were introduced in FRAYNE, MOREL and SCOTT [1962], and regular ultrafilters in KEISLER [1964]. The notion of a good ultrafilter is also due to KEISLER [1964i] where the results given here about them were first proved.

Most of the results about the set theoretic properties and the cardinality of ultraproducts can be found in FRAYNE, MOREL and SCOTT [1962]. Theorem 3.16 comes from KEISLER [1964] and the proof of theorem 3.21 is taken from CHANG [1967].

The Rabin–Keisler theorem is due to RABIN [1959] and KEISLER [1963]. The straightforward proof we have given here is due to CHANG [1965i].

CHAPTER 7

ELEMENTARY CLASSES

In this chapter we shall be interested in the relation between the semantic properties of classes of relational structures and their algebraic properties.

§1. The various sorts of elementary classes

Throughout this chapter we assume that $\mu \in \omega^{\alpha}$ is some arbitrary fixed type and that L is the appropriate language for structures of this type.

We let $\boldsymbol{M}_{\mu}$ be the class of all relational structures of type μ. We assume that all the structures we mention are in $\boldsymbol{M}_{\mu}$. If Σ is a set of sentences of L we let Σ^* be the set of all sentences deducible from Σ. We let $\boldsymbol{M}(\Sigma)$ be the collection of all models of Σ, and if $K \subseteq \boldsymbol{M}_{\mu}$ we let $Th(K)$ be the set of all sentences true in each structure in K. $Th(K)$ is called the *theory* of K. Thus

$$Th(K) = \{\sigma : \sigma \text{ is a sentence of L and for all } \mathfrak{A} \in K, \mathfrak{A} \models \sigma\},$$
$$\boldsymbol{M}(\Sigma) = \{\mathfrak{A} : \mathfrak{A} \in \boldsymbol{M}_{\mu} \text{ and for all } \sigma \in \Sigma, \mathfrak{A} \models \sigma\}.$$

Note that it follows from these definitions that $\boldsymbol{M}(\emptyset) = \boldsymbol{M}_{\mu}$ and $Th(\emptyset)$ is the set of all sentences of L.

The relationship between these notions given in the next theorem is an immediate consequence of the completeness theorem for predicate calculus.

THEOREM 1.1. *For each set Σ of sentences of* L,

$$Th(\boldsymbol{M}(\Sigma)) = \Sigma^*.$$

PROOF. If $\sigma \in \Sigma^*$ then, by the result of exercise 3.3.9, σ is true in each model of Σ, that is $\sigma \in Th(\boldsymbol{M}(\Sigma))$.

If $\sigma \notin \Sigma^*$, then not $\Sigma \vdash \sigma$ and therefore $\Sigma \cup \{\neg\sigma\}$ is consistent. Hence, by the completeness theorem, $\Sigma \cup \{\neg\sigma\}$ has a model, say $\mathfrak{A}$. $\mathfrak{A} \in \boldsymbol{M}(\Sigma)$ but σ is not true in $\mathfrak{A}$, thus $\sigma \notin Th(\boldsymbol{M}(\Sigma))$. This completes the proof.

We write $\boldsymbol{M}(\sigma)$ for $\boldsymbol{M}(\{\sigma\})$ and $Th(\mathfrak{A})$ for $Th(\{\mathfrak{A}\})$. $Th(\mathfrak{A})$ is called the *theory of* $\mathfrak{A}$.

EXERCISES

1.2. Show that for any collection $\{\sigma_i : i \in I\}$ of sentences of L,

$$\bigcap_{i \in I} \boldsymbol{M}(\sigma_i) = \boldsymbol{M}(\bigcup_{i \in I} \{\sigma_i\}).$$

1.3. Show that for any set Σ of sentences of L and any $K \subseteq \boldsymbol{M}_\mu$

$$\boldsymbol{M}(Th(\boldsymbol{M}(\Sigma))) = \boldsymbol{M}(\Sigma),$$

and

$$Th(\boldsymbol{M}(Th(K))) = Th(K).$$

1.4. Show that for any two relational structures $\mathfrak{A}, \mathfrak{B}$

$$\mathfrak{A} \equiv \mathfrak{B} \quad \text{iff} \quad Th(\mathfrak{A}) = Th(\mathfrak{B}) \quad \text{iff} \quad \mathfrak{B} \in \boldsymbol{M}(Th(\mathfrak{A})).$$

A collection K of relational structures is called an *elementary class*, if for some sentence σ of L, $K = \boldsymbol{M}(\sigma)$. If K is an elementary class we write $K \in EC$. In the terminology of ch. 5, $K \in EC$ iff the property of being a structure in K is a first order property. K is said to be *Δ-elementary*, and we write $K \in EC_\Delta$, if for some set Σ of sentences of L, $K = \boldsymbol{M}(\Sigma)$. Thus $K \in EC_\Delta$ iff the property of being a structure in K is a general first order property. By exercise 1.2, $K \in EC_\Delta$ iff for some set $\{\sigma_i : i \in I\}$ of sentences of L, $K = \bigcap_{i \in I} \boldsymbol{M}(\sigma_i)$, that is, iff K is the intersection of a set of elementary classes. We say that K is *Σ-elementary*, and we write $K \in EC_\Sigma$, if K is the union of a set of elementary classes, that is, if $K = \bigcup_{i \in I} \boldsymbol{M}(\sigma_i)$ for some set $\{\sigma_i : i \in I\}$ of sentences of L. Finally, we say that K is *$\Delta\Sigma$-elementary* and we write $K \in EC_{\Delta\Sigma}$, if K is the union of a collection of Δ-elementary classes, and we say that K is *$\Sigma\Delta$-elementary* if K is the intersection of a set of Σ-elementary classes. In this last case we write $K \in EC_{\Sigma\Delta}$.

Some relationships between these notions are given by the next lemma. If $K \subseteq \boldsymbol{M}_\mu$, then we let $cK = \boldsymbol{M}_\mu - K$.

LEMMA 1.5. *For any* $K \subseteq \boldsymbol{M}_\mu$,

(1). $K \in EC$ *iff* $cK \in EC$,

(2). $K \in EC_\Delta$ *iff* $K = \boldsymbol{M}(Th(K))$,

(3). $K \in EC_\Delta$ *iff* $cK \in EC_\Sigma$,

(4). $K \in EC_{\Delta\Sigma}$ *iff* $cK \in EC_{\Sigma\Delta}$ *iff* $cK \in EC_{\Delta\Sigma}$ *iff* $K \in EC_{\Sigma\Delta}$.

PROOF. (1). Suppose $K \in EC$, say $K = \boldsymbol{M}(\sigma)$. Then $cK = \boldsymbol{M}(\neg\sigma)$ and so $cK \in EC$. The converse is similar.

(2). If $K = \boldsymbol{M}(Th(K))$ then $K \in EC_\Delta$. Conversely if $K \in EC_\Delta$, $K = \boldsymbol{M}(\Sigma)$ for some set Σ of sentences of L. Thus $\boldsymbol{M}(Th(K)) = \boldsymbol{M}(Th(\boldsymbol{M}(\Sigma))) = \boldsymbol{M}(\Sigma) = K$, by exercise 1.3.

(3). Suppose $K \in EC_\Delta$, say $K = \boldsymbol{M}(\Sigma)$. Then

$$cK = \boldsymbol{M}_\mu - \boldsymbol{M}(\Sigma) = \boldsymbol{M}_\mu - \bigcap_{\sigma \in \Sigma} \boldsymbol{M}(\sigma) = \bigcup_{\sigma \in \Sigma} (\boldsymbol{M}_\mu - \boldsymbol{M}(\sigma)) = \bigcup_{\sigma \in \Sigma} \boldsymbol{M}(\neg\sigma),$$

and therefore $cK \in EC_\Sigma$. The converse is similar.

(4). Clearly

$$K \in EC_{\Sigma\Delta} \quad \text{iff} \quad cK \in EC_{\Delta\Sigma}$$

and

$$K \in EC_{\Delta\Sigma} \quad \text{iff} \quad cK \in EC_{\Sigma\Delta}.$$

Suppose $K \in EC_{\Sigma\Delta}$, say

$$K = \bigcap_{i \in I} K_i$$

where, for $i \in I$,

$$K_i = \bigcup_{j \in J_i} \boldsymbol{M}(\sigma_{ij}).$$

Let $J = \prod_{i \in I} J_i$, then, by the axiom of choice,

$$K = \bigcap_{i \in I} \bigcup_{j \in J_i} \boldsymbol{M}(\sigma_{ij}) = \bigcup_{f \in J} \bigcap_{i \in I} \boldsymbol{M}(\sigma_{if(i)}) \qquad [*]$$

and so $K \in EC_{\Delta\Sigma}$. Conversely if $K \in EC_{\Delta\Sigma}$, then $cK \in EC_{\Sigma\Delta}$, whence $cK \in EC_{\Delta\Sigma}$ and $K \in EC_{\Sigma\Delta}$.

EXERCISE 1.6. Verify the identity [*] in the proof above. (Hint: [*] is a set theoretic identity; it does not depend on any model theoretic properties of the sets $\boldsymbol{M}(\sigma_{ij})$.)

EXAMPLE 1.7. The collection K_G of all groups is an elementary class since, in the notation of example 5.3.11, $K_G = \boldsymbol{M}(\sigma_G)$.

Similarly all the other examples of finitely axiomatizable notions given in ch. 5 give rise to examples of elementary classes.

EXAMPLES

1.8. The collection K_0 of all commutative fields of characteristic zero is a Δ-elementary class since, in the notation of example 5.3.16, $K_0 = \boldsymbol{M}(\Delta_0)$. Theorem 5.3.17 shows that K_0 is not an elementary class.

1.9. The collection K_{fin} of all finite sets is a Σ-elementary class since $K_{\text{fin}} = \bigcup_{n \in \omega} \boldsymbol{M}(\exists!^n)$. Corollary 5.3.5 shows that K_{fin} is not an elementary class.

1.10. We shall show that the collection of infinite, finite dimensional vector spaces is a $\Delta\Sigma$-elementary class. We adopt the notation of examples 5.3.18 and 5.3.19.

Let

$$\sigma_{m,n} = \sigma_{\mathrm{VS}} \wedge \tau_{\mathrm{Dim}(m)} \wedge \exists^{\geqslant n},$$

thus $\mathfrak{A}$ is a model of $\sigma_{m,n}$ iff it is a vector space of dimension m containing at least n elements. Hence, for each m,

$$VS_m = \bigcap_{n \in \omega} \boldsymbol{M}(\sigma_{m,n})$$

is the collection of infinite vector spaces of dimension m. Therefore

$$VS = \bigcup_{m \in \omega} \bigcap_{n \in \omega} \boldsymbol{M}(\sigma_{m,n})$$

is the class of infinite, finite dimensional vector spaces. Clearly $VS \in EC_{\Delta\Sigma}$. We show that VS is not in EC_Δ or EC_Σ.

If $K = \boldsymbol{M}(\Sigma) \in EC_\Delta$, it follows from Łoš's theorem that K is closed under the formation of ultraproducts. Now for each $m \in \omega$, there is an infinite vector space of dimension m, say $\mathfrak{A}_m$. Thus $\mathfrak{A}_m$ is a model of

$$\{\sigma_{\mathrm{VS}}, \tau_{\mathrm{Dim}(m)}\} \cup \{\exists^{\geqslant n} : n \in \omega\}.$$

Let F be a non-principal ultrafilter on ω, then, using an argument that we have often used before, it can be seen that $\prod \mathfrak{A}_m/F$ is a model of

$$\{\sigma_{\mathrm{VS}}\} \cup \{\neg \tau_{\mathrm{Dim}(n)} : n \in \omega\} \cup \{\exists^{\geqslant n} : n \in \omega\}$$

and so is an infinite dimensional vector space. Therefore VS is not closed under the formation of ultraproducts and so is not in EC_Δ.

Let m be some integer and p some prime. For each $k \in \omega$ there is a field containing p^k elements and hence a vector space of dimension m containing p^{km} elements, say $\mathfrak{B}_k$. Each $\mathfrak{B}_k$ is not in VS since it is finite. But, by an argument like that used above, if F is a non-principal ultrafilter on ω, $\prod \mathfrak{B}_k/F$ is an infinite m-dimensional vector space and so is in VS. Thus cVS is also not closed under the formation of ultraproducts and therefore is not in EC_Δ It follows that VS is not in EC_Σ.

An alternative semantic characterization of $EC_{\Delta\Sigma}$ is given by the next lemma.

LEMMA 1.11 *$K \in EC_{\Delta\Sigma}$ iff K is closed under elementary equivalence, that is, if $\mathfrak{A} \in K$ and $\mathfrak{B} \equiv \mathfrak{A}$ implies $\mathfrak{B} \in K$.*

PROOF. Clearly if $K \in EC_{\Delta\Sigma}$, then K is closed under elementary equivalence. For the converse, suppose that K is closed under elementary equivalence.

If $\mathfrak{A} \in K$, by exercise 1.4, $M(Th(\mathfrak{A})) \subseteq K$, and so

$$\bigcup_{\mathfrak{A} \in K} M(Th(\mathfrak{A})) \subseteq K.$$

Clearly the converse inclusion also holds and so

$$K = \bigcup_{\mathfrak{A} \in K} M(Th(\mathfrak{A})),$$

and therefore $K \in EC_{\Delta\Sigma}$.

COROLLARY 1.12. *If $K \in EC$, $K \in EC_\Delta$ or $K \in EC_\Sigma$ then K is closed under elementary equivalence.*

An important collection of relational structures which does not come into any of the categories we have just described is given by the next result.

LEMMA 1.13. *The collection of all well-ordered sets is not even a $\Delta\Sigma$-elementary class.*

PROOF. Let $\mathfrak{A} = \langle \omega, < \rangle$ be an ordered set of order type ω, and let F be a non-principal ultrafilter on ω. Then $\mathfrak{A}$ is elementarily equivalent to $\mathfrak{A}^\omega/F$, but by theorem 6.4.3, $\mathfrak{A}^\omega/F$ is not well-ordered. Since $\mathfrak{A}$ is well-ordered, the collection of well-ordered sets is not closed under elementary equivalence and so, by lemma 1.11, is not a $\Delta\Sigma$-elementary class.

We conclude this section with some facts about the various sorts of elementary classes that can be deduced from the compactness theorem.

LEMMA 1.14. *If $\{K_i : i \in I\}$ is a collection of elementary classes, such that*

$$\bigcap_{i \in I} K_i = \emptyset,$$

then there is some finite subset I_0 of I such that

$$\bigcap_{i \in I_0} K_i = \emptyset.$$

PROOF. Suppose $K_i = M(\sigma_i)$. Since $\bigcap K_i = \emptyset$, the set of sentences $\{\sigma_i : i \in I\}$ has no model and therefore, by the compactness theorem, there is some finite subset I_0 of I such that $\{\sigma_i : i \in I_0\}$ has no model. That is,

$$\bigcap_{i \in I_0} K_i = \emptyset.$$

THEOREM 1.15. $EC = EC_\Delta \cap EC_\Sigma$.

PROOF. Trivially $EC \subseteq EC_\Delta \cap EC_\Sigma$.

Suppose $K \in EC_\Delta$ and $K \in EC_\Sigma$. Then, by lemma 1.5, K and cK are both in EC_Δ. Say

$$K = \bigcap_{i \in I} \boldsymbol{M}(\sigma_i),$$

and

$$cK = \bigcap_{j \in J} \boldsymbol{M}(\tau_j).$$

Then

$$K \cap cK = \bigcap_{i \in I} \boldsymbol{M}(\sigma_i) \cap \bigcap_{j \in J} \boldsymbol{M}(\tau_j) = \emptyset.$$

Therefore, by the previous lemma, there are finite subsets I_0, J_0 of I, J respectively, such that

$$\bigcap_{i \in I_0} \boldsymbol{M}(\sigma_i) \cap \bigcap_{j \in J_0} \boldsymbol{M}(\tau_j) = \emptyset. \qquad [*]$$

Let σ be the conjunction of the sentences $\{\sigma_i : i \in I_0\}$. Then $\bigcap_{i \in I_0} \boldsymbol{M}(\sigma_i) = \boldsymbol{M}(\sigma)$. Clearly

$$K \subseteq \boldsymbol{M}(\sigma),$$

and similarly

$$cK \subseteq \bigcap_{j \in J_0} \boldsymbol{M}(\tau_j),$$

and hence

$$c\Big(\bigcap_{j \in J_0} \boldsymbol{M}(\tau_j)\Big) \subseteq K.$$

But by [*],

$$\boldsymbol{M}(\sigma) \subseteq c\Big(\bigcap_{j \in J_0} \boldsymbol{M}(\tau_j)\Big) \subseteq K,$$

and therefore

$$K = \boldsymbol{M}(\sigma).$$

This completes the proof that $K \in EC$.

We have defined the various species of elementary classes in terms of sentences of the language L. Since these sentences only express statements about the algebraic structure of relational structures, it would seem plausible that we might be able to discover a purely algebraic characterization of these classes which makes no reference to the language L. In fact it is possible to do this. We devote the next section to the main step in this characterization.

§ 2. Keisler's ultrapower theorem

In this section we give an algebraic characterization of elementary equivalence, which has previously been defined in purely semantic terms.

Recall that we are assuming that all the relational structures we mention are of type $\mu \in \omega^{\alpha}$ and that L is the appropriate language for structures of this type. For any ordinal γ, L_γ is the language obtained from L by adding the sequence $\langle c_\xi : \xi < \gamma \rangle$ of constants.

Let β be an infinite cardinal $\geqslant \alpha$, let I be a set of cardinal β and for each $i \in I$ let $\mathfrak{A}_i$ and $\mathfrak{B}_i$ be relational structures of type μ and of cardinal at most β^+. We investigate the relationship between the following two statements:

(i). *There are uniform ultrafilters F,G on I such that the ultraproducts $\prod \mathfrak{A}_i/F$ and $\prod \mathfrak{B}_i/G$ are isomorphic.*

(ii). *There is no sentence σ of* L *such that both*

$$\{i \in I : \mathfrak{A}_i \models \sigma\}$$

and

$$\{i \in I : \mathfrak{B}_i \models \neg\sigma\}$$

are of cardinal $< \beta$.

LEMMA 2.1. *If* (i) *holds, so also does* (ii).

PROOF. Suppose F,G are uniform ultrafilters on I such that the ultraproducts $\prod \mathfrak{A}_i/F$ and $\prod \mathfrak{B}_i/G$ are isomorphic. We show that the assumption that (ii) fails leads to a contradiction.

Suppose σ is a sentence of L such that both $X = \{i \in I : \mathfrak{A}_i \models \sigma\}$ and $Y = \{i \in I : \mathfrak{B}_i \models \neg\sigma\}$ are of cardinal $< \beta$. Then, since F is uniform, $X \notin F$ and therefore, by Łoš's theorem, $\prod \mathfrak{A}_i/F \models \neg\sigma$. Similarly $\prod \mathfrak{B}_i/G \models \sigma$. This contradicts the fact that $\prod \mathfrak{A}_i/F$ is isomorphic to $\prod \mathfrak{B}_i/G$.

We devote most of the remainder of this section to showing that, assuming the generalized continuum hypothesis, (ii) implies (i). The motive for the construction we eventually use springs from the next lemma.

LEMMA 2.2. *If* (ii) *holds, there exist uniform ultrafilters F,G on I such that the ultraproducts $\prod \mathfrak{A}_i/F$ and $\prod \mathfrak{B}_i/G$ are elementarily equivalent.*

PROOF. Assume (ii) holds. For each sentence σ of L let

$$X(\sigma) = \{i \in I : \mathfrak{A}_i \models \sigma\},$$

and

$$Y(\sigma) = \{i \in I : \mathfrak{B}_i \models \sigma\}.$$

Now let

$$F_0 = \{X(\sigma) : \operatorname{card}(Y(\neg\sigma)) < \beta\}.$$

Since (ii) holds, all sets in F_0 are of cardinal β. Because $X(\sigma) \cap X(\tau) = X(\sigma \wedge \tau)$ and $Y(\neg(\sigma \wedge \tau)) = Y(\neg\sigma) \cup Y(\neg\tau)$, F_0 is closed under finite intersections.

It follows that $F_0 \cup \bar{S}_\beta(I)$ has the finite intersection property, and therefore can be extended to a (necessarily uniform) ultrafilter F on I.

Next let

$$G_0 = \{Y(\sigma) : X(\sigma) \in F\}.$$

If $Y(\tau) \in G_0$, then for some σ with $X(\sigma) \in F$, $Y(\tau) = Y(\sigma)$.

$$X(\neg\sigma) = I - X(\sigma)$$

and so $X(\neg\sigma) \notin F$. Hence, by the definition of F_0, $\mathrm{card}(Y(\neg\neg\sigma)) = \beta$, but $Y(\neg\neg\sigma) = Y(\sigma)$. This shows that all sets in G_0 are of cardinal β. It is clear that G_0 is closed under finite intersections. Therefore $G_0 \cup \bar{S}_\beta(I)$ has the finite intersection property and so can be extended to a uniform ultrafilter G on I.

If $X(\sigma) \in F$,

$$Y(\sigma) \in G_0 \subseteq G.$$

Conversely if $Y(\sigma) \in G$, $Y(\neg\sigma) \notin G$ and hence $X(\neg\sigma) \notin F$, so, since F is an ultrafilter, $X(\sigma) \in F$. Thus for each sentence σ of L,

$$\{i \in I : \mathfrak{A}_i \models \sigma\} \in F \quad \text{iff} \quad \{i \in I : \mathfrak{B}_i \models \sigma\} \in G,$$

and hence, by Łoš's theorem,

$$\prod \mathfrak{A}_i/F \models \sigma \quad \text{iff} \quad \prod \mathfrak{B}_i/G \models \sigma.$$

Therefore $\prod \mathfrak{A}_i/F$ is elementarily equivalent to $\prod \mathfrak{B}_i/G$ and the lemma is proved.

EXERCISE 2.3. Let (i)′ be the statement which results on replacing 'isomorphic' by 'elementarily equivalent' in (i). Show, without assuming that $\beta \geqslant \alpha$, or that the structures $\mathfrak{A}_i$, $\mathfrak{B}_i$ are of cardinal $\leqslant \beta^+$, that (i)′ is equivalent to (ii).

Lemma 2.2 enables us to deduce the existence of elementarily equivalent ultraproducts from (ii). We want to show that (ii) implies the existence of isomorphic ultraproducts. At this point we recall that we have previously observed that if we know that $(\mathfrak{A},a) \equiv (\mathfrak{B},b)$, where a,b are enumerations of A,B respectively, then we can at once deduce that $\mathfrak{A} \simeq \mathfrak{B}$ (theorem 4.1.12). Thus in order to show that (ii) implies the existence of isomorphic ultraproducts it is sufficient to show that (ii) implies the existence of uniform ultrafilters F,G on I, and enumerations s/F, t/G of $\prod A_i/F$ and $\prod B_i/G$, respectively, such that $(\prod \mathfrak{A}_i/F, s/F) \equiv (\prod \mathfrak{B}_i/G, t/G)$. And lemma 2.2 shows

that in order to do this it will be sufficient to show that

(iii). *There exist enumerations* $s=\langle s_\xi:\xi<\beta^+\rangle$ *and* $t=\langle t_\xi:\xi<\beta^+\rangle$ *of* $\prod A_i$ *and* $\prod B_i$ *respectively such that there is no sentence* σ *of* L_{β^+} *such that both*

$$\{i\in I:(\mathfrak{A}_i,s(i))\models\sigma\} \quad \textit{and} \quad \{i\in I:(\mathfrak{B}_i,t(i))\models\neg\sigma\},$$

are of cardinal $<\beta$, *where for* $i\in I$, $s(i)=\langle s_\xi(i):\xi<\beta^+\rangle$ *and* $t(i)=\langle t_\xi(i):\xi<\beta^+\rangle$.

A necessary condition for the existence of the enumerations s,t is obviously that card $(\prod A_i)$ and card $(\prod B_i)$ are both at most β^+. To ensure that this is the case we assume from this point onwards that the generalized continuum hypothesis holds. Then

$$\text{card}\,(\prod A_i)\leqslant\prod_{i\in I}\beta^+\leqslant(\beta^+)^\beta=\beta^+$$

and similarly for card $(\prod B_i)$. We write '(G.C.H.)' after the numbers of results that depend on the G.C.H.

If $s=\langle s_\xi:\xi<\gamma\rangle$ is a γ-termed sequence and $\delta\leqslant\gamma$, then we use the notation $s\restriction\delta$ for the δ-termed sequence $\langle s_\xi:\xi<\delta\rangle$.

LEMMA 2.4. (G.C.H.) *If* (ii) *holds, so also does* (iii).

PROOF. Assume (ii) holds. Let $a=\langle a_\xi:\xi<\beta^+\rangle$ and $b=\langle b_\xi:\xi<\beta^+\rangle$ be fixed enumerations of $\prod\mathfrak{A}_i$ and $\prod\mathfrak{B}_i$ respectively. We show that we can define by recursion a sequence $s=\langle s_\xi:\xi<\beta^+\rangle$ of elements of $\prod A_i$ and a sequence $t=\langle t_\xi:\xi<\beta^+\rangle$ of elements of $\prod B_i$ such that for all $\xi<\beta^+$:

(1). If λ is a limit ordinal, $n<\omega$, and $\xi=\lambda+2n$, then

$$s_\xi=a_{\lambda+n}$$

(2). If λ is a limit ordinal, $n<\omega$, and $\xi=\lambda+2n+1$, then

$$t_\xi=b_{\lambda+n}$$

(3). There is no sentence σ of $L_{\xi+1}$ such that both

$$\{i\in I:(\mathfrak{A}_i,s\restriction\xi+1(i))\models\sigma\} \quad \text{and} \quad \{i\in I:(\mathfrak{B}_i,t\restriction\xi+1(i))\models\neg\sigma\}$$

are of cardinal $<\beta$.

Condition (1) implies that the 'even' places in s are occupied by the elements of a and hence that s is an enumeration of $\prod A_i$. Similarly (2) implies that t is an enumeration of $\prod B_i$.

Suppose now that $\zeta<\beta^+$ and that we have defined s_ξ and t_ξ for $\xi<\zeta$ so that (1), (2) and (3) hold for $\xi<\zeta$. In the case $\zeta=0$, this is just the hypothesis that (ii) holds. We show that we can define s_ζ and t_ζ so that (1), (2) and (3) hold for all $\xi\leqslant\zeta$.

Observe first that since, if ζ is a limit ordinal any sentence of L_ζ is a sentence of $L_{\xi+1}$ for some $\xi<\zeta$, it follows that whether or not ζ is a limit ordinal, (3) holds when $\xi+1$ is replaced by ζ.

Let λ be the largest limit ordinal $\leqslant\zeta$. We consider first the case that for some $n<\omega$, $\zeta=\lambda+2n$. In this case we put

$$s_\zeta=a_{\lambda+n}$$

With this choice of s_ζ, (1) and (2) hold for all $\xi\leqslant\zeta$. It remains only to show that we can choose t_ζ so that (3) also holds for all $\xi\leqslant\zeta$. This is the crux of the proof.

Let Δ be the set of all those formulas $\phi(v_0)$ of L_ζ, containing at most the variable v_0 free, such that

$$H_\phi=\{i\in I:(\mathfrak{A}_i,s\restriction\zeta(i))\models\phi[s_\zeta(i)]\}$$

is of cardinal $<\beta$. Since $\alpha\leqslant\beta$ and $\zeta<\beta^+$, Δ is of cardinal $\leqslant\beta$.

$$\bar{H}_\phi=\{i\in I:(\mathfrak{A}_i,s\restriction\zeta(i))\models(\forall v_0)\,\phi(v_0)\}\subseteq H_\phi,$$

and therefore, if $\phi\in\Delta$, $\operatorname{card}(\bar{H}_\phi)<\beta$. Now $(\forall v_0)\phi(v_0)$ is a sentence of L_ζ and we have already observed that (3) holds with $\xi+1$ replaced by ζ. We can conclude from this that if $\phi\in\Delta$, the set

$$K_\phi=\{i\in I:(\mathfrak{B}_i,t\restriction\zeta(i))\models\neg(\forall v_0)\,\phi(v_0)\}$$

is of cardinal β.

Thus $\{K_\phi:\phi\in\Delta\}$ is a collection of not more than β sets each of cardinal β. Therefore, by lemma 6.1.20, there is a collection $\{K'_\phi:\phi\in\Delta\}$ of pairwise disjoint sets each of cardinal β such that for $\phi\in\Delta$, $K'_\phi\subseteq K_\phi$.

For each $i\in K'_\phi$ there is some $d_i\in B_i$ such that

$$(\mathfrak{B}_i,t\restriction\zeta(i))\models\neg\phi[d_i]$$

We choose $t_\zeta\in\prod B_i$ so that whenever $i\in K'_\phi$, $t_\zeta(i)$ is one such element d_i. This can be done since the sets K'_ϕ are pairwise disjoint. It remains to show that (3) holds when $\xi=\zeta$.

Let σ be a sentence of $L_{\zeta+1}$ such that the set

$$\{i\in I:(\mathfrak{A}_i,s\restriction\zeta+1(i))\models\sigma\}$$

is of cardinal $<\beta$. Since we know that (3) holds when $\xi+1$ is replaced by ζ we need only consider the case when the constant c_ζ occurs in σ, so that

$$\sigma=\phi(c_\zeta),$$

where $\phi(v_0)$ is a formula of L_ζ with at most the variable v_0 free. Thus

$$\{i \in I : (\mathfrak{A}_i, s \restriction \zeta(i)) \models \phi[s_\zeta(i)]\}$$

is of cardinal $< \beta$. Therefore $\phi \in \Delta$ and so by the choice of t_ζ whenever $i \in K'_\phi$,

$$(\mathfrak{B}_i, t \restriction \zeta) \models \neg\phi[t_\zeta(i)],$$

and so, since $\operatorname{card}(K'_\phi) = \beta$,

$$\{i \in I : (\mathfrak{B}_i, t \restriction \zeta + 1(i)) \models \neg\sigma\}$$

is of cardinal β. This shows that (3) holds for all $\xi \leqslant \zeta$.

The case when for some $n < \omega$, $\zeta = \lambda + 2n + 1$ is dealt with similarly. This completes the proof that we can find sequences s,t satisfying (1), (2) and (3) for all $\xi < \beta^+$. Since (1) is satisfied, s is an enumeration of $\prod A_i$. Since (2) is satisfied t is an enumeration of $\prod B_i$. Therefore because (3) holds we can conclude that s,t satisfy the conditions of (iii). This completes the proof of the lemma.

We can at once deduce

THEOREM 2.5. (G.C.H.) (i) *holds iff* (ii) *holds.*

PROOF. Lemma 2.1 shows that (i) implies (ii). Suppose (ii) holds, then by the previous lemma, (iii) holds. Therefore by lemma 2.2, there exist uniform ultrafilters F,G on I such that

$$\prod (\mathfrak{A}_i, s(i))/F \equiv \prod (\mathfrak{B}_i, t(i))/G \tag{1}$$

Let $s/F = \langle s_\xi/F : \xi < \beta^+ \rangle$ and $t/G = (t_\xi/G : \xi < \beta^+\rangle$. By (1),

$$(\prod \mathfrak{A}_i/F, s/F) \equiv (\prod \mathfrak{B}_i/G, t/G) \tag{2}$$

s is an enumeration of $\prod A_i$ and so s/F is an enumeration of $\prod A_i/F$. Similarly t/G is an enumeration of $\prod B_i/G$. Therefore by (2) and theorem 4.1.12,

$$\prod \mathfrak{A}_i/F \simeq \prod \mathfrak{B}_i/G.$$

This proves the theorem.

The main result of this section can be deduced at once from this last theorem.

THEOREM 2.6. (G.C.H.) KEISLER'S ULTRAPOWER THEOREM
Let β be an infinite cardinal $\geqslant \alpha$ and let $\mathfrak{A}$ and $\mathfrak{B}$ be relational structures of cardinal $\leqslant \beta^+$. $\mathfrak{A}$ is elementarily equivalent to $\mathfrak{B}$ iff there are uniform ultrafilters F,G on β such that the ultrapowers $\mathfrak{A}^\beta/F$ and $\mathfrak{B}^\beta/G$ are isomorphic.

PROOF. This is just the special case of theorem 2.5, with $I = \beta$, and for $i \in I$, $\mathfrak{A}_i = \mathfrak{A}$ and $\mathfrak{B}_i = \mathfrak{B}$.

COROLLARY 2.7. (G.C.H.) *Two relational structures* $\mathfrak{A},\mathfrak{B}$ *are elementarily equivalent iff they have isomorphic ultrapowers.*

§ 3. Characterizing elementary classes

We are now ready to set about giving an algebraic characterization of elementary classes. For each subset K of $\boldsymbol{M}_\mu$ we let

$Pow(K) = \{\mathfrak{A} \in \boldsymbol{M}_\mu$: $\mathfrak{A}$ is isomorphic to an ultrapower of some structure in $K\}$, and

$Prod(K) = \{\mathfrak{A} \in \boldsymbol{M}_\mu$: $\mathfrak{A}$ is isomorphic to an ultraproduct of elements of $K\}$.

We write $Pow(\mathfrak{A})$ for $Pow(\{\mathfrak{A}\})$ and $Prod(\mathfrak{A})$ for $Prod(\{\mathfrak{A}\})$. $Pow(K)$ and $Prod(K)$ are built up from K in a purely set theoretic fashion, that is, without any reference to the language L.

Some simple facts about *Pow* and *Prod* are contained in the following exercises.

EXERCISE 3.1. Show that:

(a). $Prod(K) \subseteq Prod(K')$ and $Pow(K) \subseteq Pow(K')$, if $K \subseteq K'$.
(b). $K \subseteq Pow(K) \subseteq Prod(K)$, for all $K \subseteq \boldsymbol{M}_\mu$.
(c). $Pow(\mathfrak{A}) = Prod(\mathfrak{A})$, for all $\mathfrak{A} \in \boldsymbol{M}_\mu$.
(d). $Pow(K) = \bigcup \{Pow(\mathfrak{A}) : \mathfrak{A} \in K\}$, if $K \subseteq \boldsymbol{M}_\mu$.
(e). $Pow(\mathfrak{A}) \subseteq \boldsymbol{M}(Th(\mathfrak{A}))$, for all $\mathfrak{A} \in \boldsymbol{M}_\mu$.
(f). $Prod(K) \subseteq \boldsymbol{M}(Th(K))$, if $K \subseteq \boldsymbol{M}_\mu$.

If $K = Pow(K)$ we say that K is *closed under isomorphism and the formation of ultrapowers.* If $K = Prod(K)$ we say that K is *closed under isomorphism and the formation of ultraproducts.*

LEMMA 3.2. *If K is closed under elementary equivalence then both K and cK are closed under isomorphism and the formation of ultrapowers.*

PROOF. This is an immediate consequence of the fact that any ultrapower of a relational structure $\mathfrak{A}$ is elementarily equivalent to $\mathfrak{A}$.

COROLLARY 3.3. *If $K \in EC_{\Delta\Sigma}$, then $K = Pow(K)$ and $cK = Pow(cK)$.*

PROOF. This is just a restatement of the lemma, using lemma 1.11.

THEOREM 3.4. *If $K \in \boldsymbol{M}_\mu$, then $K \in EC_\Delta$ iff*

(1) *K is closed under isomorphism and the formation of ultraproducts, and*

(2). *K is closed under elementary equivalence.*

PROOF. Suppose $K \in EC_\Delta$, say $K = \boldsymbol{M}(\Sigma)$. Then, trivially, K is closed under isomorphism and elementary equivalence, and, by Łoš's theorem, also under the formation of ultraproducts.

Conversely suppose (1) and (2) hold. Clearly $K \subseteq \boldsymbol{M}(Th(K))$; we show that the converse inclusion also holds.

Suppose $\mathfrak{A} \in \boldsymbol{M}(Th(K))$. Let

$$\{\sigma_i : i \in I\}$$

be the set of all sentences true in $\mathfrak{A}$. For each $i \in I$ there is some $\mathfrak{B}_i \in K$ which is a model of σ_i, or else $\neg\sigma_i$ would be in $Th(K)$ and hence true in $\mathfrak{A}$. For $j \in I$, let

$$J_j = \{i \in I : \mathfrak{B}_i \vDash \sigma_j\}.$$

It is easily seen that $\{J_j : j \in I\}$ has the finite intersection property and so can be extended to an ultrafilter F on I. Let $\mathfrak{B} = \prod \mathfrak{B}_i/F$, since K is closed under the formation of ultraproducts $\mathfrak{B} \in K$. By the construction of F each sentence σ_i is true in $\mathfrak{B}$. Therefore $\mathfrak{A} \equiv \mathfrak{B}$ and since K is closed under elementary equivalence $\mathfrak{A} \in K$. This completes the proof that $K = \boldsymbol{M}(Th(K))$ and hence that $K \in EC_\Delta$.

COROLLARY 3.5. *For any* $K \subseteq \boldsymbol{M}_\mu$,

$$K \in EC_\Delta \quad \textit{iff} \quad K \in EC_{\Delta\Sigma} \quad \textit{and} \quad K = Prod(K).$$

THEOREM 3.6. *If* $K \in \boldsymbol{M}_\mu$, *then* K *is an elementary class iff*

(1). *K and cK are closed under isomorphism and the formation of ultraproducts,*

and

(2). *K is closed under elementary equivalence.*

PROOF. It is easily seen that K is closed under elementary equivalence iff cK is. The theorem now follows immediately from theorem 3.4 and theorem 1.15.

COROLLARY 3.7. *For any* $K \in \boldsymbol{M}_\mu$,

$$K \in EC \quad \textit{iff} \quad K \in EC_{\Delta\Sigma},\ K = Prod(K) \textit{ and } cK = Prod(cK).$$

Theorems 3.4 and 3.6 contain characterizations of EC_Δ and EC which still contain the semantic notion of 'elementary equivalence'. We can eliminate this reference to the language L by using Keisler's ultrapower theorem. This theorem depends on the generalized continuum hypothesis and therefore so too do the characterizations below.

THEOREM 3.8. (G.C.H.) *For any $K \subseteq \boldsymbol{M}_\mu$, K is closed under elementary equivalence iff both K and cK are closed under isomorphism and the formation of ultrapowers.*

PROOF. One way round is lemma 3.2. For the converse suppose K and cK are closed under isomorphism and the formation of ultrapowers. Assume $\mathfrak{A} \in K$ and $\mathfrak{B} \equiv \mathfrak{A}$. Then by Keisler's ultrapower theorem, $\mathfrak{A}$ and $\mathfrak{B}$ have isomorphic ultrapowers, say $\mathfrak{A}^I/F$ and $\mathfrak{B}^I/G$. Since K is closed under the formation of ultrapowers, $\mathfrak{A}^I/F \in K$. Therefore, because K is closed under isomorphism, $\mathfrak{B}^I/G \in K$. It must therefore be the case that $\mathfrak{B} \in K$, or else, since cK is also closed under the formation of ultrapowers, we would have $\mathfrak{B}^I/G \in cK$.

It should be noted that if K is closed under isomorphism then automatically so also is cK.

COROLLARY 3.9 (G.C.H.) *For any $K \subseteq \boldsymbol{M}_\mu$,*

$$K \in EC_{\Delta\Sigma} \quad \textit{iff} \quad K = Pow(K) \textit{ and } cK = Pow(cK).$$

THEOREM 3.10. (G.C.H.) *For any $K \subseteq \boldsymbol{M}_\mu$, $K \in EC_\Delta$ iff*

(1). *K is closed under isomorphism and the formation of ultraproducts, and*

(2). *cK is closed under the formation of ultrapowers.*

PROOF. This is an immediate consequence of theorems 3.4 and 3.8,

THEOREM 3.11. (G.C.H.) *For any $K \subseteq \boldsymbol{M}_\mu$, $K \in EC$ iff both K and cK are closed under isomorphism and the formation of ultraproducts.*

PROOF. This is an immediate consequence of theorems 3.6 and 3.8.

In the next chapter we will obtain set theoretic characterizations of the various elementary classes which do not depend on the generalized continuum hypothesis.

§4. Craig's interpolation lemma

In this section we use Keisler's ultrapower theorem to deduce a classical interpolation result for predicate calculus. We need first a few new notions.

Suppose that $\mathfrak{A} = \langle A, \{R_\xi : \xi < \alpha\}\rangle$ is a relational structure of type $\mu \in \omega^\alpha$. An *expansion* of $\mathfrak{A}$ is a relational structure of the form

$$\mathfrak{A}' = \langle A, \{R'_\xi : \xi < \alpha + \beta\}\rangle,$$

where, for $\xi < \alpha$, $R'_\xi = R_\xi$. If $\mathfrak{A}'$ is an expansion of $\mathfrak{A}$ then we say that $\mathfrak{A}$ is a *reduct* of $\mathfrak{A}'$. If $\mathfrak{A}$ is of type $\mu \in \omega^\alpha$ and $\mathfrak{A}'$ is of type $\mu' \in \omega^{\alpha+\beta}$, then $\mu \subseteq \mu'$.

If L is the appropriate language for structures of type μ then the appropriate language for talking about $\mathfrak{A}'$ is obtained by adding a new set $\{P_\xi : \alpha \leqslant \xi < \alpha + \beta\}$ of predicate letters to L to form a new language L', called an *expansion* of L. We say that $\mathfrak{A}$ is the *reduction of* $\mathfrak{A}'$ *to* L. Thus $\mathfrak{A}$ is obtained from $\mathfrak{A}'$ by dropping all the relations corresponding to predicate letters of L' not in L.

If $K \subseteq \boldsymbol{M}_\mu$, we say that K is a *pseudo-elementary class* if there is an expansion L' of L and an elementary class with respect to the language L', say K', such that K is the class of reductions to the language L of the elements of K'. If K is a pseudo-elementary class we write $K \in PC$.

EXAMPLE 4.1. Let L be the language appropriate for groups and let L' be the expansion of L obtained by adding a new binary predicate letter P. Let σ_{COG} be the conjunction of the sentence σ_{CG} of example 5.3.12 and the following sentences

$$(\forall x)(\forall y)(\forall z)[S(x,y,z) \wedge P(e,x) \wedge P(e,y) \rightarrow P(e,z)],$$
$$(\forall x)(\forall y)[S(x,y,e) \rightarrow \neg(P(e,x) \wedge P(e,y))],$$
$$(\forall x)(x = e \vee P(e,x) \vee (\forall y)[S(x,y,e) \rightarrow P(e,y)]),$$

then a realization $\mathfrak{A}'$ of L' is a model of σ_{COG} iff it is an ordered group. Thus the collection of ordered groups is an elementary class with respect to the language L'. Hence the class of *orderable groups* is a pseudo-elementary class with respect to the language L, since it is the collection of reductions of all ordered groups to the language L.

We begin with two obvious remarks.

LEMMA 4.2. *For each $i \in I$, let $\mathfrak{A}_i$ be the reduction of $\mathfrak{A}_i'$ to the language* L *and let F be an ultrafilter on I. Then $\prod \mathfrak{A}_i/F$ is the reduction of $\prod \mathfrak{A}_i'/F$ to the language* L.

LEMMA 4.3. *If $K \in PC$, K is closed under isomorphism and the formation of ultraproducts, i.e. $K = Prod(K)$.*

Our aim is to prove a separation theorem for pseudo-elementary classes, namely that disjoint elements of PC are separated by an elementary class. We do this by a series of lemmas.

LEMMA 4.4. *For any $K \subseteq \boldsymbol{M}_\mu$, $Prod(K)$ contains a model of each consistent extension of $Th(K)$.*

PROOF. Let Δ be any consistent extension of $Th(K)$ and let $\Sigma = \{\sigma_i : i \in I\}$ be the set of all finite conjunctions of sentences of Δ. Σ is also consistent.

For each $i \in I$, there is some $\mathfrak{A}_i \in K$ which is a model of σ_i or else $\neg\sigma_i$

would be in $Th(K)$ and hence in Σ, contradicting the fact that Σ is consistent.

For each $j \in I$, let

$$J_j = \{i \in I : \mathfrak{A}_i \vDash \sigma_j\}.$$

Each J_j is non-empty and since Σ is closed under finite conjunctions, $\{J_j : j \in I\}$ has the finite intersection property and so can be extended to an ultrafilter F on I.

By construction, for each $\sigma \in \Sigma$,

$$\{i \in I : \mathfrak{A}_i \vDash \sigma\} \in F$$

and therefore, using Łoš's theorem, $\prod \mathfrak{A}_i/F$ is a model of Σ. Since $\prod \mathfrak{A}_i/F \in Prod(K)$ the lemma is proved.

LEMMA 4.5. *For any $K, K' \subseteq \boldsymbol{M}_\mu$, $Th(K) \cup Th(K')$ is consistent iff some ultraproduct of elements of K is elementarily equivalent to some ultraproduct of elements of K'.*

PROOF. Suppose $\mathfrak{A} \in Prod(K)$, $\mathfrak{B} \in Prod(K')$ and $\mathfrak{A} \equiv \mathfrak{B}$. $\mathfrak{A}$ is a model of $Th(K)$ and $\mathfrak{B}$ is a model of $Th(K')$. Therefore $\mathfrak{A}$ is a model of $Th(K) \cup Th(K')$, which is therefore consistent.

Conversely suppose $Th(K) \cup Th(K')$ is consistent. By Lindenbaum's theorem (theorem 9.1.19) $Th(K) \cup Th(K')$ has a complete consistent extension Σ. That is

$$Th(K) \cup Th(K') \subseteq \Sigma$$

and for any sentence σ of L, either σ or $\neg\sigma$ is in Σ.

By the previous lemma, there exist $\mathfrak{A} \in Prod(K)$ and $\mathfrak{B} \in Prod(K')$ which are both models of Σ. Since Σ is complete it follows that $\mathfrak{A} \equiv \mathfrak{B}$.

LEMMA 4.6. (G.C.H.) *For any $K, K' \subseteq \boldsymbol{M}_\mu$, $Th(K) \cup Th(K')$ is consistent iff $Prod(K) \cap Prod(K') \neq \emptyset$.*

PROOF. One way round is an immediate consequence of the previous lemma.

For the converse, suppose $Th(K) \cup Th(K')$ is consistent. Then by lemma 4.5, there exist $\mathfrak{A} \in Prod(K)$ and $\mathfrak{B} \in Prod(K')$ with $\mathfrak{A} \equiv \mathfrak{B}$. By Keisler's ultrapower theorem, there is an ultrapower $\mathfrak{A}^I/F$ of $\mathfrak{A}$ and an isomorphic ultrapower $\mathfrak{B}^I/G$ of $\mathfrak{B}$. By exercise 6.2.7, any ultrapower of an ultraproduct of elements of K is itself an ultraproduct of elements of K, so $\mathfrak{A}^I/F \in Prod(K)$. Similarly $\mathfrak{B}^I/G \in Prod(K')$. Since $\mathfrak{A}^I/F \simeq \mathfrak{B}^I/G$, $Prod(K) \cap Prod(K')$ is not empty.

LEMMA 4.7. (G.C.H.) *If $K = Prod(K)$, $K' = Prod(K')$ and $K \cap K' = \emptyset$, then there is some $N \in EC$ such that $K \subseteq N$ and $N \cap K' = \emptyset$.*

PROOF. If $Th(K) \cup Th(K')$ were consistent, by lemma 4.6,

$$Prod(K) \cap Prod(K')$$

and hence also $K \cap K'$ would not be empty. Therefore $Th(K) \cup Th(K')$ must be inconsistent. Hence there is some finite subset Σ_0 of $Th(K) \cup Th(K')$ which is inconsistent and therefore has no model.

Let σ_0 be the conjunction of the finite number of sentences in $\Sigma_0 \cap Th(K)$ if this set is not empty, otherwise let σ_0 be any universally valid sentence, e.g. $(\forall v_0)(v_0 = v_0)$. Let $N = \boldsymbol{M}(\sigma_0)$; $N \in EC$. Since $\sigma_0 \in Th(K)$, $K \subseteq N$ and since $Th(K') \cup \{\sigma_0\}$ is inconsistent, $N \cap K' = \emptyset$.

LEMMA 4.8. (G.C.H.) *If $K, K' \in PC$ and $K \cap K' = \emptyset$, there is some $N \in EC$ such that $K \subseteq N$ and $N \cap K' = \emptyset$.*

PROOF. This is an immediate consequence of lemmas 4.3 and 4.7.

From lemma 4.8 we are able to deduce the main result of this section.

THEOREM 4.9. (G.C.H.) CRAIG'S INTERPOLATION LEMMA
If σ, τ are sentences of predicate calculus such that

$$\vdash \sigma \rightarrow \tau,$$

then there is some sentence θ containing only predicate letters common to σ and τ such that both

$$\vdash \sigma \rightarrow \theta \quad \textit{and} \quad \vdash \theta \rightarrow \tau.$$

PROOF. Let L′ be the language containing all the predicate letters in σ and τ and let L be the language containing just the predicate letters common to σ and τ. L′ is an expansion of L.

Let K be the class of all reductions to L of the realizations of L′ which are models of σ, and let K' be the class of all reductions to L of those realizations of L′ which are models of $\neg\tau$. $K, K' \in PC$.

Since $\vdash \sigma \rightarrow \tau$, $\{\sigma, \neg\tau\}$ is inconsistent and therefore $K \cap K' = \emptyset$.* Therefore by the previous lemma, there is some $N \in EC$ such that $K \subseteq N$ and $N \cap K' = \emptyset$. Say $N = \boldsymbol{M}(\theta)$, where θ is some sentence of L.

Because $K \subseteq N$, there is no model of $\{\sigma, \neg\theta\}$ and therefore by the completeness theorem $\vdash \sigma \rightarrow \theta$. Similarly since $N \cap K' = \emptyset$, $\vdash \theta \rightarrow \tau$.

This completes the proof.

The original proof of this result was a syntactic one which gives a little more

* Puzzled readers should read the note on pp. 159–160.

information about θ than that given above (see CRAIG [1957]). More important is the fact that the original proof made no use of the G.C.H. However its use in the proof above can be eliminated by the following consideration. By using an appropriate Gödel numbering of the symbols of L', theorem 4.9 can be restated as a sentence of first order arithmetic. It follows from Gödel's proof of the relative consistency of the G.C.H. (GÖDEL [1940]) that any such sentence which is provable using the G.C.H. can be proved without using this hypothesis, or even the axiom of choice (for a discussion of this see, e.g. KREISEL [1956]). Therefore we can conclude from theorem 4.9. that Craig's lemma is true without assuming the G.C.H.

We give another semantic proof of Craig's lemma, very similar to that above, but not using the G.C.H. in the next chapter, where there is a further discussion of this theorem.

§5. EC as a Boolean algebra

Although we have been careful not to call M_μ, EC or the elements of EC *sets*, we have talked about them as though they were perfectly good sets, writing such things as $M(\sigma) \in EC$. But, as is well known, too free and easy a use of naive set theoretic ideas leads to contradictions (e.g. Russell's paradox). The contradictions can (we hope) be avoided by using, for example, Von Neumann–Bernays–Gödel set theory. We remark, therefore, that in general M_μ is a proper class in the sense of GÖDEL [1940] and similarly for most $K \subseteq M_\mu$. Therefore no set theoretic meaning can really be attached to expressions such as '$K \in EC$'. We can get over this difficulty by working instead of with M_μ, with a *subset* of M_μ, which although sufficiently small to be a set, is still roomy enough to allow the standard model theoretic constructions (formation of ultrapowers, productions of models of consistent sets of sentences etc.) to be carried out within it. We do not pursue this point any further here, but we leave it to the reader who is worried about our naive talk of EC, to translate everything we do into statements about real sets.

Since for any sentences σ, τ of L,

$$M(\sigma) \cap M(\tau) = M(\sigma \wedge \tau) \quad \text{and} \quad M_\mu - M(\sigma) = M(\neg \sigma),$$

it is easily seen that EC is a Boolean subalgebra of the Boolean algebra $\langle \mathscr{P}(M_\mu), \subseteq \rangle$. Since EC is closed under finite intersections it may be taken as a base for a topology (the *elementary* topology) on M_μ. Clearly EC_Δ and EC_Σ are the families of closed sets and open sets, respectively, in this topology. By theorem 1.15, $EC = EC_\Delta \cap EC_\Sigma$ is the family of clopen sets in this topo-

logy. Furthermore it follows from the compactness theorem that if a subset Z of EC (i.e. a family of closed sets) has the finite intersection property then $\bigcap Z$ is non-empty (see lemma 1.14) and this means, of course, that $\boldsymbol{M}_\mu$ is compact with the elementary topology. This explains how the compactness theorem got its name.

The remarks above show that $\boldsymbol{M}_\mu$ with the elementary topology is a compact space with a base of clopen sets so that if we could show that $\boldsymbol{M}_\mu$ was a Hausdorff space we would have established that it was a Boolean space. Unfortunately the elementary topology does not satisfy any of the standard separation conditions, indeed it is not even a T_0 space. (If $\mathfrak{A},\mathfrak{B}\in\boldsymbol{M}_\mu$ and $\mathfrak{A}\equiv\mathfrak{B}$, every open set containing $\mathfrak{A}$ contains $\mathfrak{B}$ and vice versa.) This drawback at first sight would seem to thwart our design of applying Stone's results to the subject. With inexhaustible ingenuity however (this is J.L.B.'s remark, not mine—A.B.S.), we overcome this difficulty by applying Kolmogoroff's well known device for obtaining a T_0 space from an arbitrary space. This runs as follows: let Y be an arbitrary topological space and let the closure of any subset A of Y be denoted by clA. We define the relation $\sim$ on Y by $x\sim y$ iff $cl\{x\}=cl\{y\}$, for all $x,y\in Y$. Clearly $\sim$ is an equivalence relation and the quotient space $Y^*=Y/\sim$ with the resulting topology is a T_0 space (the *Kolmogoroff space* of Y). The canonical map $\pi: Y\to Y^*$ which sends each element of Y to its equivalence class in Y^* is continuous. Also $\pi^{-1}\pi[U]=U$ for any open set U in Y. We now have

THEOREM 5.1. *The Kolmogoroff space $\boldsymbol{M}_\mu^*$ of $\boldsymbol{M}_\mu$ is a Boolean space and the map $\sigma: EC\to\mathscr{P}(\boldsymbol{M}_\mu^*)$ defined by $\sigma(K)=\pi[K]$, for each $K\in EC$, where $\pi:\boldsymbol{M}_\mu\to\boldsymbol{M}_\mu^*$ is the canonical map, is an isomorphism of EC onto the characteristic algebra EC^* of $\boldsymbol{M}_\mu^*$.*

PROOF. The space $\boldsymbol{M}_\mu^*$ as the continuous image (under π) of the compact space $\boldsymbol{M}_\mu$ is certainly compact, and it is a T_0 space by construction. Moreover the base EC for $\boldsymbol{M}_\mu$ is mapped by π onto a base $EC^*=\{\pi[K]: K\in EC\}$ for $\boldsymbol{M}_\mu^*$ and each member of this base is clopen. But a T_0 space with a base of clopen sets is Hausdorff and so $\boldsymbol{M}_\mu^*$ is Hausdorff and hence a Boolean space.

We leave it to the reader to verify that σ is indeed an isomorphism.

According to this theorem $\boldsymbol{M}_\mu^*$ may be regarded, up to homeomorphism, as the Stone space of EC.

For any member $\mathfrak{A}\in\boldsymbol{M}_\mu$ it can easily be seen that $cl\{\mathfrak{A}\}=\{\mathfrak{B}:\mathfrak{A}\equiv\mathfrak{B}\}$, so that the space $\boldsymbol{M}_\mu^*$ may equally well be regarded as the quotient space of

M_μ modulo elementary equivalence. We now prove an interesting representation theorem which illustrates the duality between the semantic and the syntactic aspects of the theory.

THEOREM 5.2. *EC is isomorphic to the Lindenbaum algebra* $\mathscr{A}(\mathrm{L})$ *of the sentences of the language* L.

PROOF. Clearly as we have already remarked, for any sentences σ,τ of L,

$$M(\sigma \wedge \tau) = M(\sigma) \cap M(\tau),$$

$$M(\sigma \vee \tau) = M(\sigma) \cup M(\tau)$$

and

$$M(\neg\sigma) = M_\mu - M(\sigma).$$

Furthermore, the sentences σ,τ are provably equivalent iff $M(\sigma)=M(\tau)$, so that the map $M^*:\mathscr{A}(\mathrm{L})\to EC$ defined by $M^*(|\sigma|)=M(\sigma)$, where $|\sigma|$ is the equivalence class in $\mathscr{A}(\mathrm{L})$ containing σ, is well defined and in fact is an isomorphism.

COROLLARY 5.3. *The space* M_μ^* *is homeomorphic to the Stone space of* $\mathscr{A}(\mathrm{L})$.

PROOF. We have just shown that their characteristic algebras are isomorphic.

§ 6. Historical and bibliographical remarks

The notion of an elementary class is due to Tarski (see e.g. TARSKI [1950]). Keisler's ultrapower theorem is taken from KEISLER [1961]. Theorems 3.4 and 3.6 can be found in FRAYNE, MOREL and SCOTT [1962] and theorems 3.8 and 3.11 in KEISLER [1961]. Craig's interpolation theorem was first proved by CRAIG [1957]. The proof we have given is due to KEISLER [1961], the remark about eliminating the use of the G.C.H. in this proof is based on a remark of Kreisel's (in KREISEL [1956]). It was TARSKI [1950] who first pointed out the topological significance of elementary classes and the compactness theorem.

Note. These remarks will clarify our assertion in the proof of theorem 4.9 that since $\{\sigma,\neg\tau\}$ is inconsistent, $K\cap K'=\emptyset$.

Let $\{P_i':1\leqslant i\leqslant m\}$ be the set of predicate letters occurring in σ but not in τ and let $\{P_j'':1\leqslant j\leqslant n\}$ be the set of predicate letters occurring in τ but not in σ. Suppose $K\cap K'\neq\emptyset$, say $\mathfrak{A}\in K\cap K'$, with

$$\mathfrak{A}=\langle A,\{R_\xi:\xi<\alpha\}\rangle.$$

Since $\mathfrak{A} \in K$, there is an expansion $\mathfrak{A}'$ of $\mathfrak{A}$ which is a model of σ, say

$$\mathfrak{A}' = \langle A, \{R_\xi : \xi < \alpha\} \cup \{R'_i : 1 \leqslant i \leqslant m\} \cup \{R''_j : 1 \leqslant j \leqslant n\} \rangle.$$

Since $\mathfrak{A} \in K'$, there is an expansion $\mathfrak{A}''$ of $\mathfrak{A}$ which is a model of $\neg\tau$, say

$$\mathfrak{A}'' = \langle A, \{R_\xi : \xi < \alpha\} \cup \{S'_i : 1 \leqslant i \leqslant m\} \cup \{S''_j : 1 \leqslant j \leqslant n\} \rangle.$$

But then

$$\mathfrak{A}^* = \langle A, \{R_\xi : \xi < \alpha\} \cup \{R'_i : 1 \leqslant i \leqslant m\} \cup \{S''_j : 1 \leqslant j \leqslant n\} \rangle$$

is a model of both σ and $\neg\tau$. This contradicts the fact that $\{\sigma, \neg\tau\}$ is inconsistent and hence shows that $K \cap K' = \emptyset$.

CHAPTER 8

ULTRALIMITS

In the previous chapter we showed that if two relational structures are elementarily equivalent, they have isomorphic ultrapowers, and we used this result to obtain an algebraic characterization of elementary classes. In order to obtain this result we had to assume the generalized continuum hypothesis. In this chapter we describe the ultralimit construction, due to S. B. Kochen, which enables us to achieve a similar result not depending on this hypothesis.

§1. Two basic lemmas

Once again, for simplicity, throughout most of this chapter we deal only with relational structures which consist of a non-empty set and a single binary relation on this set. We assume that such structures are of type μ, that L is the appropriate language for structures of this type and that for any ordinal ξ, L_ξ is the language obtained from L by adding the sequence $\langle c_\zeta : \zeta < \xi \rangle$ of constants.

We begin with two basic lemmas.

LEMMA 1.1. FRAYNE'S LEMMA

If $\mathfrak{A},\mathfrak{B}$ *are relational structures,* $\mathfrak{A} \equiv \mathfrak{B}$ *iff there is an ultrafilter pair* $\langle I,F \rangle$ *such that* $\mathfrak{B}$ *is elementarily embeddable in* $\mathfrak{A}^I/F$.

PROOF. If $\mathfrak{B}$ is elementarily embeddable in $\mathfrak{A}^I/F$, $\mathfrak{B} \equiv \mathfrak{A}$, and therefore, because $\mathfrak{A} \equiv \mathfrak{A}^I/F$, it follows that $\mathfrak{A} \equiv \mathfrak{B}$.

Conversely, suppose that $\mathfrak{A} \equiv \mathfrak{B}$. Let $b = \langle b_\xi : \xi < \gamma \rangle$ be an enumeration of B. We consider the extended system $(\mathfrak{B},b)$ and the extended language L_γ appropriate for structures of this type.

Let S be the set of all sentences of L_γ and let

$$I = \{\sigma \in S : (\mathfrak{B},b) \models \sigma\}.$$

Suppose $\sigma \in I$. In general σ will contain a finite number of the constants of L_γ, say $c_{\xi_1},\ldots,c_{\xi_k}$. We let σ' be the formula of L obtained from σ by substituting k variables, which do not occur in σ, say $v_{i_1},\ldots,v_{i_k}$, for these constants. Since $(\mathfrak{B},b) \models \sigma$, $\mathfrak{B} \models \theta$, where

$$\theta = (\exists v_{i_1})\ldots(\exists v_{i_k})\, \sigma' .$$

Now $\mathfrak{A} \equiv \mathfrak{B}$ and θ is a sentence of L. Therefore $\mathfrak{A} \models \theta$. It follows that there is some sequence $a_\sigma = \langle a_{\sigma\xi} : \xi < \gamma \rangle$ of elements of A such that $(\mathfrak{A},a_\sigma) \models \sigma$. Thus we can find such a sequence a_σ for each sentence σ in I.

Let

$$J_\sigma = \{\tau \in I : (\mathfrak{A},a_\tau) \models \sigma\} .$$

It is easily seen that the family $J = \{J_\sigma : \sigma \in I\}$ has the finite intersection property and so can be extended to an ultrafilter F on I. We show that $\mathfrak{B}$ can be elementarily embedded in $\mathfrak{A}^I/F$.

We define the map $h : \mathfrak{B} \to \mathfrak{A}^I/F$ as follows. If $b_\xi \in B$, we put

$$h(b_\xi) = f_\xi/F ,$$

where $f_\xi \in A^I$ is the function such that for each $\sigma \in I$

$$f_\xi(\sigma) = a_{\sigma\xi} .$$

Suppose ϕ is a formula of L and $d \in B^\omega$ is such that $\mathfrak{B} \models_d \phi$; we shall complete the proof by showing that $\mathfrak{A}^I/F \models_{h(d)} \phi$.

Let ϕ' be the sentence of L_γ obtained from ϕ as follows: for each $k < \omega$, if d_k is b_ξ replace all free occurrences of v_k in ϕ by c_ξ. Clearly

$$\mathfrak{B} \models_d \phi \quad \text{iff} \quad (\mathfrak{B},b) \models \phi' .$$

But

$$(\mathfrak{B},b) \models \phi' \quad \text{iff} \quad \phi' \in I \quad \text{iff} \quad J_{\phi'} \in F .$$

That is

$$\text{iff} \quad \{\psi \in I : (\mathfrak{A},a_\psi) \models \phi'\} \in F ,$$

and by Łoś's theorem this holds

$$\text{iff} \quad (\mathfrak{A}^I/F, f/F) \models \phi' ,$$

where $f_\psi = a_\psi$ for $\psi \in I$.

But, by the definition, $h(b)=f/F$, and so we have shown that

$$(\mathfrak{B},b)\models\phi' \quad \text{iff} \quad (\mathfrak{A}^I/F,h(b))\models\phi',$$

from which it at once follows that

$$\mathfrak{B}\models_a\phi \quad \text{iff} \quad \mathfrak{A}^I/F\models_{h(d)}\phi,$$

and hence h is an elementary embedding.

We note in passing the following simple consequence of Frayne's lemma.

COROLLARY 1.2. *If $\mathfrak{A},\mathfrak{B}$ are elementarily equivalent, finite relational structures, they are isomorphic.*

PROOF. By Frayne's lemma, there is an ultrafilter pair $\langle I,F\rangle$ such that both $\mathfrak{A}$ and $\mathfrak{B}$ are elementarily embeddable in $\mathfrak{A}^I/F$. Since A is finite, $\mathrm{card}(A)=\mathrm{card}(A^I/F)$ by lemma 6.3.6, and since $\mathfrak{A}\equiv\mathfrak{B}$, $\mathrm{card}(A)=\mathrm{card}(B)$. Hence the elementary embeddings of $\mathfrak{A},\mathfrak{B}$ into $\mathfrak{A}^I/F$ must be onto maps and so isomorphisms. That is

$$\mathfrak{A}\simeq\mathfrak{A}^I/F\simeq\mathfrak{B}.$$

LEMMA 1.3. SCOTT'S LEMMA

A map $g:\mathfrak{A}\to\mathfrak{B}$ is an elementary embedding iff there is an ultrafilter pair $\langle I,F\rangle$ and an elementary embedding $h:\mathfrak{B}\to\mathfrak{A}^I/F$ such that diagram 1.a *is commutative, where d is the canonical embedding,*

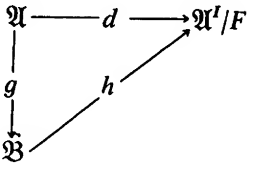

Diagram 1.a

i.e. such that for all $a\in A$, $h(g(a))=d(a)$.

PROOF. Suppose first that there is such an ultrafilter pair and such an embedding $h:\mathfrak{B}\to\mathfrak{A}^I/F$. If $\mathfrak{A}\models_x\phi$, then, since d is an elementary embedding, $\mathfrak{A}^I/F\models_{d(x)}\phi$, that is $\mathfrak{A}^I/F\models_{h(g(x))}\phi$. Therefore, because h is an elementary embedding $\mathfrak{B}\models_{g(x)}\phi$. This shows that g is an elementary embedding.

Conversely, suppose that $g:\mathfrak{A}\to\mathfrak{B}$ is an elementary embedding. As in the proof of the previous lemma, we suppose that the sequence $b=\langle b_\xi:\xi<\gamma\rangle$ is an enumeration of B and we let S and I be defined as before.

Let a_+ be some fixed element of A. If $b_\xi \in g[A]$ let $b'_\xi = b_\xi$ and otherwise let $b'_\xi = g(a_+)$. Then the sequence

$$b' = \langle b'_\xi : \xi < \gamma \rangle \in (g[A])^\gamma .$$

For each $\xi < \gamma$ let a_ξ be $g^{-1}(b_\xi)$. Since $\mathfrak{A} \simeq g[\mathfrak{A}] \prec \mathfrak{B}$, by theorem 4.1.10,

$$(\mathfrak{A}, a) \equiv (\mathfrak{B}, b') \qquad [*]$$

where a is the sequence $\langle a_\xi : \xi < \gamma \rangle$.

Now suppose $\phi \in I$. Let $c_{\xi_0}, \ldots, c_{\xi_k}$ be those constants occurring in ϕ such that the corresponding elements of B, $b_{\xi_0}, \ldots, b_{\xi_k}$ are not in $g[A]$. Let ϕ' be the formula of L_γ obtained from ϕ by replacing, for $i \leqslant k$, each c_{ξ_i} by a variable v_{j_i}, not occurring in ϕ. Since $(\mathfrak{B}, b) \models \phi$,

$$(\mathfrak{B}, b') \models (\exists v_{j_0}) \ldots (\exists v_{j_k}) \phi'$$

and so by [*],

$$(\mathfrak{A}, a) \models (\exists v_{j_0}) \ldots (\exists v_{j_k}) \phi' .$$

Hence there is some sequence

$$a_\phi = \langle a_{\phi\xi} : \xi < \gamma \rangle$$

of elements of A such that

$$(\mathfrak{A}, a_\phi) \models \phi$$

where, if ξ is not one of the ξ_i with $i \leqslant k$, $a_{\phi\xi} = a_\xi$.

Thus we can find such a sequence a_ϕ for each $\phi \in I$. Just as in Frayne's lemma, if

$$J_\phi = \{\psi \in I : (\mathfrak{A}, a_\psi) \models \phi\},$$

then $J = \{J_\phi : \phi \in I\}$ has the finite intersection property and so can be extended to an ultrafilter F on I. We define the map $h : \mathfrak{B} \to \mathfrak{A}^I/F$ as in the previous lemma. Again h is an elementary embedding, so it remains only to show that diagram 1.a is commutative, that is, that $d = h \cdot g$.

Suppose $a' \in A$, then $g(a') \in B$, say $g(a') = b_\xi$. Hence

$$h \cdot g(a') = f_\xi / F,$$

where $f_\xi(\phi) = a_{\phi\xi}$.

$$b_\xi \in g[A]$$

and so

$$f_\xi(\phi) = a_{\phi\xi} = a_\xi = g^{-1}(b_\xi) = a' .$$

Thus $f_\xi / F = d(a')$ and we have shown that $h \cdot g = d$.

§ 2. Ultralimits

We have already seen that a relational structure $\mathfrak{A}$ can be embedded in any of its ultrapowers, so that we can regard $\mathfrak{A}^I/F$ as an elementary extension of $\mathfrak{A}$. Similarly the ultrapower $(\mathfrak{A}^I/F)^J/G$ can be regarded as an extension of $\mathfrak{A}^I/F$. So by reiterating the ultrapower construction we obtain essentially an increasing elementary chain of relational structures. If we repeat the process $\aleph_0$ times and take the union of the chain the resulting structure is called an *ultralimit* of $\mathfrak{A}$. We now give a more formal definition which takes into account that $\mathfrak{A}^I/F$ is an elementary extension of $\mathfrak{A}$ only up to isomorphism. It is this fact which causes the complication in the definition that follows.

By an *ultrafilter sequence* we mean a sequence

$$\langle\langle I_n,F_n\rangle : n<\omega\rangle, \qquad [*]$$

where for each $n<\omega$, $\langle I_n,F_n\rangle$ is an ultrafilter pair.

Now suppose $\mathfrak{A}=\langle A,R\rangle$ is a relational structure and that U is an ultrafilter sequence of the form [*]. We define the sequence $\langle\mathfrak{A}_n : n<\omega\rangle$ of relational structures as follows:

$$\mathfrak{A}_0=\mathfrak{A}=\langle A_0,R_0\rangle, \quad \text{say},$$

and for each $n\in\omega$,

$$\mathfrak{A}_{n+1}=\mathfrak{A}_n^{I_n}/F_n=\langle A_{n+1},R_{n+1}\rangle, \text{ say.}$$

For each $n\in\omega$, let d_n be the canonical embedding $d_n:\mathfrak{A}_n\to\mathfrak{A}_{n+1}$. Let $d_n^n:\mathfrak{A}_n\to\mathfrak{A}_n$ be the identity map, and for $m\leqslant n$ let d_m^n be the composition

$$d_m^n=d_{n-1}\cdot d_{n-2}\cdot\cdots\cdot d_{m+1}\cdot d_m.$$

It is clear that d_m^n is an elementary embedding of $\mathfrak{A}_m$ into $\mathfrak{A}_n$.

Clearly, also, if $k\leqslant m\leqslant n$,

$$d_k^n=d_m^n\cdot d_k^m.$$

Let

$$A'=\bigcup_{n\in\omega} A_n.$$

We define the relation $\sim$ on A' as follows:

If $a_m\in A_m$ and $a_n\in A_n$,

$$a_m\sim a_n \quad \text{iff} \quad \text{for some } p\in\omega, \quad d_m^p(a_m)=d_n^p(a_n).$$

Obviously in this case $p\geqslant m,n$.

LEMMA 2.1. *$\sim$ is an equivalence relation on A'.*

PROOF. Since

$$d_n^n(a_n)=a_n=d_n^n(a_n),$$

$a_n\sim a_n$ and so $\sim$ is reflexive. $\sim$ is obviously symmetric. We show that $\sim$ is also a transitive relation.

Suppose $a_i\in A_i$, $a_j\in A_j$, $a_k\in A_k$, $a_i\sim a_j$ and $a_j\sim a_k$. Then for some $m\geqslant i,j$ and $n\geqslant j,k$

$$d_i^m(a_i)=d_j^m(a_j) \tag{1}$$

and

$$d_j^n(a_j)=d_k^n(a_k). \tag{2}$$

Let $p\in\omega$ be greater than m and n, then

$$\begin{aligned} d_i^p(a_i)&=d_m^p\cdot d_i^m(a_i)=d_m^p\cdot d_j^m(a_j), \quad \text{by (1)}\\ &=d_j^p(a_j)\\ &=d_n^p\cdot d_j^n(a_j)=d_n^p\cdot d_k^n(a_k), \quad \text{by (2)}\\ &=d_k^p(a_k) \end{aligned}$$

and therefore $a_i\sim a_k$. Diagram 2.a illustrates the situation.

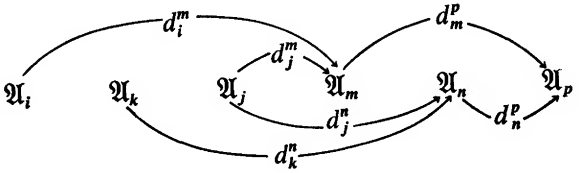

Diagram 2.a

This completes the proof.

We let A_ω be the set of equivalence classes of elements of A' under the relation $\sim$.

LEMMA 2.2. *If $\{a_i: i\leqslant k\}$ is a finite subset of A_ω, there is some $n\in\omega$ such that for $i\leqslant k$ there is an element $a_{ni}\in A_n$ which is in the equivalence class a_i.*

PROOF. For each $i\leqslant k$, there is some $n_i\in\omega$ and an $a_{n_i}\in A_{n_i}$ such that $a_{n_i}\in a_i$. Let $n\in\omega$ be greater than any of the n_i, and let

$$a_{ni}=d_{n_i}^n(a_{n_i}).$$

$a_{ni}\in A_n$ and since

$$d_{n_i}^n(a_{n_i})=a_{ni}=d_n^n(a_{ni})$$

$a_{n_i}\sim a_{ni}$ and therefore $a_{ni}\in a_i$. This completes the proof.

LEMMA 2.3. *Suppose* $m,n\in\omega$, $a_m,b_m\in A_m$, $a_n,b_n\in A_n$, $a_m\sim a_n$ *and* $b_m\sim b_n$, *then*

$$\langle a_m,b_m\rangle\in R_m \quad \textit{iff} \quad \langle a_n,b_n\rangle\in R_n.$$

PROOF. There exist $p,q\geqslant m,n$ such that

$$d^p_m(a_m)=d^p_n(a_n) \quad \text{and} \quad d^q_m(b_m)=d^q_n(b_n).$$

Let $r\in\omega$ be greater than both p and q, then clearly,

$$d^r_m(a_m)=d^r_n(a_n) \quad \text{and} \quad d^r_m(b_m)=d^r_n(b_n).$$

d^r_m and d^r_n are elementary embeddings and hence

$$\begin{aligned}\langle a_m,b_m\rangle\in R_m \quad &\text{iff} \quad \langle d^r_m(a_m),d^r_m(b_m)\rangle\in R_r\\ &\text{iff} \quad \langle d^r_n(a_n),d^r_n(b_n)\rangle\in R_r\\ &\text{iff} \quad \langle a_n,b_n\rangle\in R_n.\end{aligned}$$

This completes the proof.

We now define the relation R_ω on A_ω as follows. Suppose $a,b\in A_\omega$. By lemma 2.2 there is some $n\in\omega$ such that there is an element $a_n\in A_n\cap a$ and an element $b_n\in A_n\cap b$. Now we put

$$\langle a,b\rangle\in R_\omega \quad \text{iff} \quad \langle a_n,b_n\rangle\in R_n.$$

By lemma 2.3 this definition is independent of the choice of n and of a_n and b_n and so is a sound definition.

The relational structure

$$\mathfrak{A}_\omega=\langle A_\omega,R_\omega\rangle$$

is called the *ultralimit* of $\mathfrak{A}$ with respect to the ultrafilter sequence U. There is nothing in our notation to show with respect to which ultrafilter sequence a particular ultralimit is defined, but this will not matter as this will always be obvious from the context.

For each $n\in\omega$, let $d^n:\mathfrak{A}_n\to\mathfrak{A}_\omega$ be the map which sends each element of A_n onto its equivalence class in A_ω.

LEMMA 2.4. *Each* d^n *is an elementary embedding.*

PROOF. The proof of this lemma is by induction on the number of logical symbols in formulas and is essentially the same as the proof of theorem 4.2.1 on the union of elementary chains. We leave the details to the reader.

The map d^n is called the *projection* of $\mathfrak{A}_n$ into $\mathfrak{A}_\omega$.

We now come to the main result about ultralimits.

THEOREM 2.5. KOCHEN'S ULTRALIMIT THEOREM
$\mathfrak{A}$ is elementarily equivalent to $\mathfrak{B}$ iff $\mathfrak{A}$ and $\mathfrak{B}$ have isomorphic ultralimits.

PROOF. Let $\mathfrak{A}_\omega, \mathfrak{B}_\omega$ be isomorphic ultralimits of $\mathfrak{A}$ and $\mathfrak{B}$ respectively. By the previous lemma $\mathfrak{A}$ is elementarily embeddable in $\mathfrak{A}_\omega$ and so $\mathfrak{A} \equiv \mathfrak{A}_\omega$. Similarly $\mathfrak{B} \equiv \mathfrak{B}_\omega$. Since $\mathfrak{A}_\omega \simeq \mathfrak{B}_\omega$, it follows that $\mathfrak{A} \equiv \mathfrak{B}$. This proves one half of the theorem.

For the converse assume $\mathfrak{A} \equiv \mathfrak{B}$. Let $\mathfrak{A}_0 = \mathfrak{A}$ and $\mathfrak{B}_0 = \mathfrak{B}$. By Frayne's lemma there is an ultrafilter pair $\langle I_0, F_0 \rangle$ and an elementary embedding h_0 of $\mathfrak{B}_0$ into $\mathfrak{A}_0^{I_0}/F_0 = \mathfrak{A}_1$, say. Hence, by Scott's lemma, there is an ultrafilter pair $\langle J_0, G_0 \rangle$ such that diagram 2.b is commutative, where $\mathfrak{B}_1 = \mathfrak{B}_0^{J_0}/G_0$, g_0 is an elementary embedding, and d_0, e_0 are the canonical embeddings.

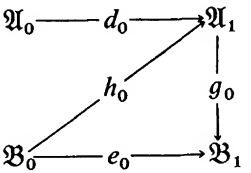

Diagram 2.b

Clearly, by an obvious inductive argument, using Scott's lemma repeatedly, the diagram above can be extended to the infinite commutative diagram 2.c, where the maps g_n, h_n are elementary embeddings, the maps d_n, e_n are canonical embeddings, $\mathfrak{A}_{n+1} = \mathfrak{A}_n^{I_n}/F_n$ for some ultrafilter pair $\langle I_n, F_n \rangle$ and $\mathfrak{B}_{n+1} = \mathfrak{B}_n^{J_n}/G_n$ for some ultrafilter pair $\langle J_n, G_n \rangle$.

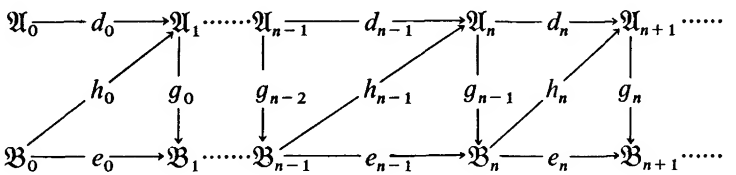

Diagram 2.c

Let $\mathfrak{A}_\omega$ be the ultralimit of $\mathfrak{A}$ with respect to the ultrafilter sequence $\langle \langle I_n, F_n \rangle : n \in \omega \rangle$ and let $\mathfrak{B}_\omega$ be the ultralimit of $\mathfrak{B}$ with respect to the ultrafilter sequence $\langle \langle J_n, G_n \rangle : n \in \omega \rangle$. We will show that $\mathfrak{A}_\omega$ is isomorphic to $\mathfrak{B}_\omega$.

For each $n\in\omega$, let d^n be the projection of $\mathfrak{A}_n$ into $\mathfrak{A}_\omega$ and let e^n be the projection of $\mathfrak{B}_n$ into $\mathfrak{B}_\omega$. We define the map $g_\omega:\mathfrak{A}_\omega\to\mathfrak{B}_\omega$ as follows:

If $a\in\mathfrak{A}_\omega$ then there is some $n\in\omega$, with some $a_n\in A_n\cap a$. We let

$$g_\omega(a)=e^n g_{n-1}(a_n).$$

Clearly g_ω is well defined.

Suppose $a,a'\in A_\omega$. Then by lemma 2.2 there is some $n\in\omega$, with some $a_n\in A_n\cap a$ and some $a'_n\in A_n\cap a'$. Then

$$\begin{aligned}\langle a,a'\rangle\in R \quad &\text{iff} \quad \langle a_n,a'_n\rangle\in R_n, \text{ by the definition of } R_\omega,\\ &\text{iff} \quad \langle g_{n-1}(a_n),g_{n-1}(a'_n)\rangle\in S_n, \text{ since } g_{n-1} \text{ is an embedding},\\ &\text{iff} \quad \langle e^n g_{n-1}(a_n),e^n g_{n-1}(a'_n)\rangle\in S_\omega, \text{ since } e^n \text{ is an embedding},\\ &\text{iff} \quad \langle g_\omega(a),g_\omega(a')\rangle\in S_\omega.\end{aligned}$$

In the same way it follows that $a=a'$ iff $g_\omega(a)=g_\omega(a')$. Therefore g_ω is an isomorphism. To complete the proof it remains only to show that g_ω maps A_ω *onto* B_ω.

Suppose $b\in B_\omega$. Then for some $n\in\omega$ there is some $b_n\in B_n\cap b$. Let $a=d^{n+1}h_n(b_n)$. Since $h_n(b_n)\in A_{n+1}\cap d^{n+1}h_n(b_n)$,

$$\begin{aligned}g_\omega(a)&=e^{n+1}g_n h_n(b_n)\\ &=e^{n+1}e_n(b_n), \text{ because 2.c is commutative},\\ &=e^n(b_n)\\ &=b.\end{aligned}$$

This shows that g_ω maps A_ω onto B_ω. Therefore g_ω is an isomorphism between $\mathfrak{A}_\omega$ and $\mathfrak{B}_\omega$. This completes the proof of the theorem.

The ultralimit construction is much more complicated than the ultrapower construction. In return for the price of greater complication we obtain the benefit that, while Keisler's ultrapower theorem depends on the G.C.H., the analogous theorem for ultralimits, Kochen's ultralimit theorem, depends on no such assumption. We exploit this result in the next section.

§3. Elementary classes and ultralimits

In §3 of ch. 7 we obtained set theoretic characterizations of EC and EC_Δ by applying Keisler's ultrapower theorem to theorems 7.3.4 and 7.3.6 to eliminate the reference to elementary equivalence. These characterizations depended on the generalized continuum hypothesis. We can now eliminate the references to elementary equivalence by making use of Kochen's ultra-

limit theorem. In this way we obtain characterizations of EC and EC_Δ which do not depend on the G.C.H.

THEOREM 3.1. *If $K \subseteq \boldsymbol{M}_\mu$, then K is closed under elementary equivalence iff both K and cK are closed under isomorphism and the formation of ultralimits.*

PROOF. One way round is trivial. Suppose then that both K and cK are closed under isomorphism and the formation of ultralimits. Let $\mathfrak{A}$ be an element of K and suppose $\mathfrak{B} \equiv \mathfrak{A}$. By Kochen's ultralimit theorem $\mathfrak{A}$ and $\mathfrak{B}$ have isomorphic ultralimits $\mathfrak{A}_\omega$ and $\mathfrak{B}_\omega$. K is closed under the formation of ultralimits and so $\mathfrak{A}_\omega \in K$. $\mathfrak{A}_\omega \simeq \mathfrak{B}_\omega$ and so $\mathfrak{B}_\omega \in K$. Hence $\mathfrak{B}$ must be in K or else, since cK is also closed under the formation of ultralimits, we would have $\mathfrak{B}_\omega \in cK$.

THEOREM 3.2. *For any $K \subseteq \boldsymbol{M}_\mu$, $K \in EC_\Delta$ iff K is closed under isomorphism, the formation of ultraproducts and ultralimits and cK is closed under the formation of ultralimits.*

PROOF. This is an immediate consequence of theorem 7.3.4 and theorem 3.1.

THEOREM 3.3. *For any $K \subseteq \boldsymbol{M}_\mu$, $K \in EC$ iff both K and cK are closed under isomorphism and the formation of ultraproducts and ultralimits.*

PROOF. This is an immediate consequence of theorems 7.3.6 and theorem 3.1.

§4. Ultralimits and Craig's interpolation lemma

Not surprisingly we can use Kochen's ultralimit theorem to eliminate the use of the generalized continuum hypothesis in the proof of Craig's interpolation lemma that we gave in ch. 7.

For any $K \subseteq \boldsymbol{M}_\mu$, let

$$Ult(K) = \{\mathfrak{A} \in \boldsymbol{M}_\mu : \mathfrak{A} \text{ is isomorphic to an ultralimit of some element of } K\}$$

LEMMA 4.1. *If $K \in PC$, then $K = Ult(Prod(K))$.*

PROOF. Trivially, $K \subseteq Ult(Prod(K))$. Suppose $\mathfrak{B} \in Ult(Prod(K))$. Then for some collection $\{\mathfrak{A}_i : i \in I\}$ of elements of K,

$$\mathfrak{B} \simeq (\prod \mathfrak{A}_i/F)_\omega .$$

Since $K \in PC$, there is some extension L′ of L and some sentence σ of L′ such that for $i \in I$, $\mathfrak{A}_i$ is the reduction to L of some model $\mathfrak{A}_i'$ of σ.

Since each $\mathfrak{A}_i'$ is a model of σ, it is easily seen that $(\prod \mathfrak{A}_i'/F)_\omega$ is also a model of σ. Also it can be seen that $(\prod \mathfrak{A}_i/F)_\omega$ is the reduction to L of $(\prod \mathfrak{A}_i'/F)_\omega$. Therefore $\mathfrak{B} \in K$. This completes the proof.

LEMMA 4.2. *For any* $K,K' \subseteq \boldsymbol{M}_\mu$, $Th(K) \cup Th(K')$ *is consistent iff*

$$Ult(Prod(K)) \cap Ult(Prod(K')) \neq \emptyset.$$

PROOF. Clearly if

$$\mathfrak{A} \in Ult(Prod(K)) \cap Ult(Prod(K')),$$

then $\mathfrak{A}$ is a model of $Th(K) \cup Th(K')$ which is therefore consistent.

Conversely suppose $Th(K) \cup Th(K')$ is consistent. Then by theorem 7.4.5. there exists $\mathfrak{A} \in Prod(K)$ and $\mathfrak{B} \in Prod(K')$ with $\mathfrak{A} \equiv \mathfrak{B}$. By Kochen's ultralimit theorem therefore, there exist ultralimits $\mathfrak{A}_\omega, \mathfrak{B}_\omega$ of $\mathfrak{A}, \mathfrak{B}$ respectively, which are isomorphic. Therefore

$$\mathfrak{A}_\omega \in Ult(Prod(K)) \cap Ult(Prod(K')).$$

THEOREM 4.3. *If* $K,K' \in PC$ *and* $K \cap K' = \emptyset$, *then there is some* $N \in EC$ *such that* $K \subseteq N$ *and* $N \cap K' = \emptyset$.

PROOF. If $K,K' \in PC$ and $K \cap K' = \emptyset$, then by lemma 4.1,

$$Ult(Prod(K)) \cap Ult(Prod(K')) = \emptyset$$

and so, by lemma 4.2, $Th(K) \cup Th(K')$ is inconsistent. The result now follows exactly as in the proof of theorem 7.4.7.

Note that theorem 4.3 is exactly the same as lemma 7.4.8 except that it no longer depends on the generalized continuum hypothesis. Therefore just as in ch. 7, §4, we can immediately deduce

THEOREM 4.4. CRAIG'S INTERPOLATION LEMMA
If σ,τ *are sentences of predicate calculus such that*

$$\vdash \sigma \rightarrow \tau$$

then there is some sentence θ *containing only predicate letters common to* σ *and* τ *such that both*

$$\vdash \sigma \rightarrow \theta \quad \textit{and} \quad \vdash \theta \rightarrow \tau.$$

This time however the proof does not depend on the G.C.H. It should be noticed, though, that we have used the axiom of choice in the proof, since

it involves the ultraproduct construction. Craig's original syntactic proof did not use the axiom of choice. The considerations we mentioned in ch. 7 that enabled us to eliminate the reference to the G.C.H. in our first proof, also enable us to eliminate reference to the axiom of choice in the second proof, if this is thought desirable.

COROLLARY 4.5. *Let* $\phi(v_1,\ldots,v_k)$, $\psi(v_1,\ldots,v_k)$ *be formulas of L such that*

$$\vdash (\forall v_1)\ldots(\forall v_k)[\phi(v_1,\ldots,v_k)\rightarrow\psi(v_1,\ldots,v_k)].$$

Then there is a formula $\theta(v_1,\ldots,v_k)$ *containing only predicate letters common to* ϕ *and* ψ *such that*

$$\vdash (\forall v_1)\ldots(\forall v_k)[\phi(v_1,\ldots,v_k)\rightarrow\theta(v_1,\ldots,v_k)],$$

and

$$\vdash (\forall v_1)\ldots(\forall v_k)[\theta(v_1,\ldots,v_k)\rightarrow\psi(v_1,\ldots,v_k)].$$

PROOF. We form the language L_{k+1} by adding the constants $c_1,\ldots,c_k$. By our hypothesis,

$$\vdash \phi(c_1,\ldots,c_k)\rightarrow\psi(c_1,\ldots,c_k)$$

and hence, by Craig's lemma, there is some sentence $\theta(c_1,\ldots,c_k)$ containing only predicate letters occurring in ϕ and ψ such that

$$\vdash \phi(c_1,\ldots,c_k)\rightarrow\theta(c_1,\ldots,c_k)$$

and

$$\vdash \theta(c_1,\ldots,c_k)\rightarrow\psi(c_1,\ldots,c_k).$$

The result now follows at once.

It should be noted that we can only assert that the free variables of θ occur among $v_1,\ldots,v_k$ not that they include all these variables. A syntactic proof would yield more information about which variables occur free in θ.

Let Σ be a set of sentences which contains the predicate letter P which is of degree k, and possibly other predicate letters. Let L be the language whose predicate letters are precisely those in Σ.

We say that P is *explicitly definable* with respect to Σ if there is a formula $\phi(v_1,\ldots,v_k)$ of L, not containing P, and whose free variables are among $v_1,\ldots,v_k$, such that

$$\Sigma\vdash (\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\leftrightarrow\phi(v_1,\ldots,v_k)].$$

We say that P is *implicitly definable* with respect to Σ if Σ determines the relation P uniquely, in the following sense. Let P' be a new k-ary predicate

letter not occurring in Σ, and let Σ' be the set of sentences obtained from Σ by replacing all occurrences of P by P'. Then Σ determines P uniquely if whenever P satisfies Σ and P' satisfies Σ' we can deduce that P and P' determine the same relation. Or, to be formally precise, P is *implicitly definable* with respect to Σ if

$$\Sigma \cup \Sigma' \vdash (\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\leftrightarrow P'(v_1,\ldots,v_k)].$$

One can give a simple model-theoretic interpretation of the notion of implicit definability. Let L′ be the language obtained from L by suppressing P; if $\mathfrak{A}$ is a realization of L′ and R is a k-ary relation over A, let $(\mathfrak{A},R)$ be the realization of L obtained by adjoining the relation R to $\mathfrak{A}$. Then it is easy to see (using the completeness theorem) that P is implicitly definable with respect to Σ iff whenever $\mathfrak{A}$ is a realization of L′ and R,S are k-ary relations over A such that $(\mathfrak{A},R)$ and $(\mathfrak{A},S)$ are both models of Σ, then $R=S$.

THEOREM 4.6. BETH'S THEOREM
P is explicitly definable with respect to Σ iff P is implicitly definable with respect to Σ.

PROOF. One way round is obvious. For the converse, suppose that P is implicitly definable with respect to Σ and hence that

$$\Sigma \cup \Sigma' \vdash (\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\leftrightarrow P'(v_1,\ldots,v_k)].$$

By the finiteness theorem there are finite subsets Σ_0,Σ_0' of Σ and Σ' respectively, such that

$$\Sigma_0 \cup \Sigma_0' \vdash (\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\leftrightarrow P'(v_1,\ldots,v_k)].$$

Let σ_0 be the conjunction of the sentences in Σ_0 and let σ_0' be the conjunction of the sentences in Σ_0'. By the deduction theorem

$$\vdash (\sigma_0 \wedge \sigma_0')\rightarrow(\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\leftrightarrow P'(v_1,\ldots,v_k)],$$

whence, clearly,

$$\vdash (\forall v_1)\ldots(\forall v_k)[(\sigma_0 \wedge P(v_1,\ldots,v_k))\rightarrow(\sigma_0'\rightarrow P'(v_1,\ldots,v_k))].$$

Hence, by the corollary to Craig's lemma, there is a formula $\phi(v_1,\ldots,v_k)$ containing only predicate letters common to $\sigma_0 \wedge P(v_1,\ldots,v_k)$ and $\sigma_0'\rightarrow P'(v_1,\ldots,v_k)$, and hence only predicate letters occurring in Σ other than P, such that

$$\vdash (\forall v_1)\ldots(\forall v_k)[(\sigma_0 \wedge P(v_1,\ldots,v_k))\rightarrow\phi(v_1,\ldots,v_k)], \tag{1}$$

and

$$\vdash (\forall v_1)\ldots(\forall v_k)[\phi(v_1,\ldots,v_k)\rightarrow(\sigma_0'\rightarrow P'(v_1,\ldots,v_k))]. \tag{2}$$

By (1)

$$\sigma_0 \vdash (\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\rightarrow\phi(v_1,\ldots,v_k)], \tag{3}$$

and by (2)

$$\sigma_0' \vdash (\forall v_1)\ldots(\forall v_k)[\phi(v_1,\ldots,v_k)\rightarrow P'(v_1,\ldots,v_k)]. \tag{4}$$

Replacing P' by P in (4) we get

$$\sigma_0 \vdash (\forall v_1)\ldots(\forall v_k)[\phi(v_1,\ldots,v_k)\rightarrow P(v_1,\ldots,v_k)]. \tag{5}$$

It follows at once from (3) and (5) that

$$\Sigma \vdash (\forall v_1)\ldots(\forall v_k)[P(v_1,\ldots,v_k)\leftrightarrow\phi(v_1,\ldots,v_k)],$$

and thus P is implicitly definable with respect to Σ.

For some applications of these theorems see for example KREISEL and KRIVINE [1967], ch. 6.

§ 5. Historical and bibliographical remarks

Frayne's lemma is due to T. Frayne and Scott's lemma to D. Scott. Proofs can be found in FRAYNE, MOREL and SCOTT [1962]. Corollary 1.2 is a weak version of a result of TARSKI [1954] who proved that two finite relational structures are isomorphic if any universal sentence holding in one also holds in the other. We leave the proof of this as an exercise for the reader. The notion of an ultralimit is due to KOCHEN [1962] where the results of §§2 and 3 can be found. Beth's theorem was first proved in BETH [1953].

CHAPTER 9

COMPLETENESS AND MODEL COMPLETENESS

In this chapter we shall be more concerned with particular mathematical structures than we have been up to now. We shall see how important results about such structures can be established by model theoretic methods.

§1. Completeness

We will be dealing in this chapter with structures of various different types. In each case we let L be the appropriate language for structures of the type we are considering. By a *theory* we mean a consistent set of sentences of L. By 'property' we will mean 'first order property'. We assume throughout that L is countable.

A question that naturally arises is that of the extent to which the properties of the models of some theory are determined by the sentences of the theory. One answer is given by the completeness theorem for predicate calculus. We have already seen (theorem 7.1.1) that one consequence of this theorem is that $Th(M(\Sigma))=\Sigma^*$, or, in other words,

THEOREM 1.1. *Let Σ be a theory and P a (first order) property, then every model of Σ has the property P iff the sentence which determines P is a consequence of Σ.*

Although this result determines the properties shared by all the models of a given theory it gives us almost no information about the extent to which the properties of a particular model of Σ are determined by Σ. In general the models of a given theory can be very different, for example, there are all sorts of groups, finite and infinite, commutative and non-commutative etc.. We are going to be interested in theories which determine their models more precisely than the group axioms do, and in how we can test a theory to see how precisely it determines its models.

The most extreme case that arises is when all models of a theory are isomorphic; such a theory is said to be *categorical.* For an example of a

categorical theory consider $\exists!^n$, the sentence described in exercise 5.3.2. The theory $\{\exists!^n\}$ is categorical since any two sets of the same cardinal are isomorphic (as sets).

EXERCISE 1.2. Find another example of a categorical theory.

Two models of a theory cannot be isomorphic unless they have the same cardinal. It follows from the Löwenheim–Skolem theorem that if a theory has an infinite model then it has a model of more than one cardinal and so cannot be categorical. Thus very few theories are categorical. However if we impose a less stringent condition and call a theory $\mathfrak{m}$-*categorical* if all models of the theory of cardinal $\mathfrak{m}$ are isomorphic then we find that it is met by a number of interesting theories. We list some of these now.

EXAMPLE 1.3. DENSELY ORDERED SETS WITHOUT FIRST OR LAST ELEMENT Let σ_{DOS} be the sentence described in example 5.3.10 and let

$$\Sigma=\{\sigma_{\text{DOS}},(\forall v_0)[(\exists v_1)\,P(v_0,v_1)\wedge(\exists v_1)\,P(v_1,v_0)]\}.$$

Σ is the theory of densely ordered sets without first or last element. A well known theorem of CANTOR [1895] says that this theory is $\aleph_0$-categorical. We give a proof of this result now.

Let $\mathfrak{A}=\langle A,<\rangle$ and $\mathfrak{B}=\langle B,<\rangle$ be two countable models of Σ. Thus each is a densely ordered set without first or last element. Let $\langle a_n:n<\omega\rangle$ be an enumeration of A and let $\langle b_n:n<\omega\rangle$ be an enumeration of B. We define the sequences $\langle c_n:n<\omega\rangle$ and $\langle d_n:n<\omega\rangle$ by recursion as follows:

$$c_0=a_0 \quad \text{and} \quad d_0=b_0.$$

Now suppose $0<k<\omega$; there are two cases to consider:

(i) $k=2m$ is even. In this case we put

$$c_k=a_m.$$

If for some $j<k$, $c_k=c_j$, then we put $d_k=d_j$. Otherwise we let d_k be some element of B which bears the same order relations to the elements of

$$\{d_0,\ldots,d_{k-1}\}$$

as does c_k to the elements of

$$\{c_0,\ldots,c_{k-1}\},$$

that is, for each $j<k$, if $c_k<c_j$ then $d_k<d_j$ and if $c_j<c_k$ then $d_j<d_k$. Since B is densely ordered by $<$ without first or last element such an element d_k can always be found.

(ii) $k=2m+1$ is odd. In this case we put

$$d_k=b_m.$$

If for some $j<k$, $d_k=d_j$, then we put $c_k=c_j$. Otherwise we let c_k be some element of A which bears the same order relations to the elements of

$$\{c_0,\ldots,c_{k-1}\}$$

as does d_k to the elements of

$$\{d_0,\ldots,d_{k-1}\}.$$

Again such a c_k can always be found.
Thus our recursive definition is completed.

We define the map $h:A\to B$, as follows:

$$h(c_n)=d_n, \quad \text{for each } n<\omega.$$

It is clear from our construction that h is an order isomorphism between $\{c_n:n\in\omega\}$ and $\{d_n:n\in\omega\}$. Also it is clear from our construction that

$$\{c_n: n\in\omega\}=\{a_n: n\in\omega\}=A$$

and

$$\{d_n: n\in\omega\}=\{b_n: n\in\omega\}=B.$$

Thus h is an order isomorphism of A onto B. This completes the proof.

EXERCISE 1.4. Show that the theories of densely ordered sets
(1). with first and last elements,
(2). with first but without last element,
(3). with last but without first element,
are all $\aleph_0$-categorical.

EXAMPLE 1.5. ALGEBRAICALLY CLOSED FIELDS OF SPECIFIED CHARACTERISTIC
It has been shown by STEINITZ [1910] that for any uncountable cardinal $\mathfrak{m}$ and any p, all algebraically closed fields of cardinal $\mathfrak{m}$ and characteristic p, are isomorphic. Thus, using the notation of examples 5.3.16 and 5.3.20, each of the theories

$$\Sigma_p=\{\sigma_F^p\}\cup\{\tau_n:2\leqslant n\leqslant\omega\}, \quad \text{for } p \text{ prime},$$

and

$$\Sigma_0=\Delta_0\cup\{\tau_n:2\leqslant n\leqslant\omega\}$$

is $\mathfrak{m}$-categorical, for each uncountable cardinal $\mathfrak{m}$.

EXAMPLE 1.6. INFINITE SETS
Clearly the theory $\{\exists^{\geqslant n}: n \in \omega\}$ is $\mathfrak{m}$-categorical for each infinite cardinal $\mathfrak{m}$, since all sets of the same cardinal are isomorphic as sets.

It is possible to show that the theory of densely ordered sets is not $\mathfrak{m}$-categorical for any uncountable $\mathfrak{m}$, and that the theories of example 1.5 are not $\aleph_0$-categorical. Example 1.6 provides an example of a theory which is $\mathfrak{m}$-categorical for all infinite cardinals $\mathfrak{m}$, and the theory of groups that of a theory which is not $\mathfrak{m}$-categorical for any infinite $\mathfrak{m}$. It is natural to enquire whether these are the only possibilities. MORLEY [1964] has shown that a theory which is $\mathfrak{m}$-categorical for *some* uncountable $\mathfrak{m}$ is $\mathfrak{m}$-categorical for *any* uncountable $\mathfrak{m}$. Thus these are the only possibilities. Morley's proof is too difficult to be given here.

We give some more examples of theories which are $\mathfrak{m}$-categorical for some $\mathfrak{m}$ later in this section.

We have seen that a theory is said to be categorical if all its models are isomorphic. Another relation which might hold between all models of a given theory is that of elementary equivalence. We say that a theory is *complete* if all its models are elementarily equivalent. Thus the models of a complete theory cannot be distinguished by sentences of the language L since they all have the same first order properties. An important property of complete theories, which is often used to define them, is given by the next result.

THEOREM 1.7. *A theory Σ is complete iff for any sentence σ of* L, *either $\Sigma \vdash \sigma$ or $\Sigma \vdash \neg\sigma$.*

PROOF. Suppose Σ is complete. Then for any sentence σ, either σ or $\neg\sigma$ holds in all models of Σ. Hence by theorem 1.1 either σ or $\neg\sigma$ is a consequence of Σ.

Conversely suppose Σ is not complete. Then there are models $\mathfrak{A},\mathfrak{B}$ of Σ such that $\mathfrak{A} \not\equiv \mathfrak{B}$. Hence for some sentence σ, $\mathfrak{A} \models \sigma$ but $\mathfrak{B} \models \neg\sigma$. Thus, using theorem 1.1, neither σ nor $\neg\sigma$ is a consequence of Σ.

EXERCISE 1.8. If $\mathfrak{A}$ is a relational structure and $\Sigma = Th(\mathfrak{A})$ show that Σ is complete.

The importance of completeness is twofold. First, if we know that a theory is complete and that some model of it has a certain first order property then we immediately know that all its models have the same property. Thus we can investigate the properties of all the models by looking at just one particular model which we can choose to be as simple as possible.

Second, readers familiar with the elements of recursive function theory will notice that *complete, recursively axiomatizable theories are decidable.* This second fact about completeness is very useful. To see why it is so, observe that the theorems of a recursively axiomatizable theory (that is, the consequences of the axioms) form a recursively enumerable set, as also do the refutable sentences of the theory (the negations of the theorems). But if a theory is complete each sentence is either a theorem or is refutable. Hence the set of theorems and its complement are both recursively enumerable and so they are both recursive sets. So the theory is decidable.

It will be seen that all the theories that we show to be complete are recursively axiomatizable, and so are decidable.

The next theorem provides a useful test for completeness.

THEOREM 1.9. VAUGHT'S TEST

Let Σ be a theory with no finite models and which is $\mathfrak{m}$-categorical for some infinite cardinal $\mathfrak{m}$, then Σ is complete.

PROOF. Suppose Σ is not complete. Then for some sentence σ neither σ nor $\neg\sigma$ is a consequence of Σ. Therefore $\Sigma \cup \{\sigma\}$ and $\Sigma \cup \{\neg\sigma\}$ are both consistent and so have models. Since Σ has no finite models, these models are infinite. Hence, by the Löwenheim–Skolem theorem, $\Sigma \cup \{\sigma\}$ and $\Sigma \cup \{\neg\sigma\}$ have models of cardinal $\mathfrak{m}$. Since σ holds in one of these models but not in the other, these models of Σ are not isomorphic. This contradicts that Σ is $\mathfrak{m}$-categorical and this contradiction proves that Σ is complete.

The applications of this theorem depend on various algebraic results. We will not give proofs of these results, but just give references to where proofs may be found. The first three applications are consequences of results we have already mentioned.

COROLLARY 1.10. *The theory of densely ordered sets without first or last elements is complete.*

This is a consequence of the result of example 1.3. It gives us another proof that the rationals, considered as an ordered set, are elementarily equivalent to the reals, also considered as an ordered set. Similarly all the theories mentioned in exercise 1.5 are complete.

COROLLARY 1.11. *The theory of an algebraically closed field of specified characteristic is complete.*

This is a consequence of the result mentioned in example 1.5.

COROLLARY 1.12. *The theory of equality for infinite sets is complete.*

This is a consequence of example 1.6.

COROLLARY 1.13. *The theory of atomless Boolean algebras is complete.*

PROOF. By exercise 1.7.3 and corollary 1.7.7 the theory of atomless Boolean algebras has only infinite models and is $\aleph_0$-categorical. Therefore, by Vaught's test, the theory of such algebras is complete.

EXERCISE 1.14. Write down a set of axioms for the theory of atomless Boolean algebras.

COROLLARY 1.15. *The theory of infinite commutative groups in which all non-zero elements are of the same prime order is complete.*

PROOF. Let G be a commutative group of cardinal $\mathfrak{m}$, whose non-zero elements are all of order p. Let X be a set of generators of the group G. X must be of cardinal $\mathfrak{m}$ since all elements of G are of order p. It is easy to see that G is isomorphic to the direct sum of the cyclic groups generated by the elements of X individually. These are all copies of the cyclic group of order p.

These remarks show that all commutative groups, whose non-zero elements are all of order p, of cardinal $\mathfrak{m}$ are isomorphic to the direct sum of $\mathfrak{m}$ copies of the cyclic group of order p. It follows that the theory of such groups is $\mathfrak{m}$-categorical for each infinite $\mathfrak{m}$. The result now follows from Vaught's test.

COROLLARY 1.16. *The theory of infinite commutative divisible, torsion free groups is complete.*

PROOF. A group is *torsion free* if all its non-zero elements are of infinite order. It is *divisible* if for each element a of the group and any integer n, there is some element b of the group with $n \cdot b = a$, where $n \cdot b$ is the sum $b+b+\cdots+b$, with n terms.

To write down a set of axioms for commutative, divisible, torsion free groups, we make use of the abbreviation $S_n(x,y)$ introduced in example 5.3.16. Recall that $S_n(x,y)$ 'says' $y = n \cdot x$.

A relational structure $\mathfrak{A}$ is a divisible, torsion free, commutative group iff it is a model of Δ_{DTF} where

$$\Delta_{\mathrm{DTF}} = \{\sigma_{\mathrm{CG}}\} \cup \{(\forall x)[x = 0 \vee \neg S_n(x,0)] : n \in \omega\} \cup \{(\forall x)(\exists y)\, S_n(y,x) : n \in \omega\}.$$

The theory Δ_{DTF} is not $\aleph_0$-categorical, since both the real rationals and the complex rationals are countable, commutative, divisible, torsion free groups

but they are not isomorphic as additive groups. However it is known that all commutative, divisible, torsion free groups are isomorphic to direct sums of the additive group of rationals (this follows from theorem 4 of KAPLANSKY [1954], page 10) and hence Δ_{DTF} is $\mathfrak{m}$-categorical for uncountable $\mathfrak{m}$. (So, for example, the real numbers and the complex numbers are isomorphic as additive groups.) Again we can now use Vaught's test to deduce the result we want, namely that Δ_{DTF} is complete.

COROLLARY 1.17. *For each integer n and prime p, the theory of n-dimensional vector spaces over an algebraically closed field of characteristic p is complete.*

PROOF. Using the notation of example 5.3.19, let

$$\Sigma_n = \{\sigma_{\text{VS}}, \tau_{\text{Dim}(n)}\}.$$

Σ_n is the theory of vector spaces of dimension n. Clearly from the theory Σ_{CF}^p of algebraically closed fields of characteristic p we can derive a set $\Sigma_{\text{CF}}^{p\prime}$ of sentences which say that the scalars of a vector space form such a field. Thus

$$\Sigma_{n,p} = \Sigma_n \cup \Sigma_{\text{CF}}^{p\prime}$$

is the theory of n-dimensional vector spaces over algebraically closed fields of characteristic p. Let $\mathfrak{A}$ and $\mathfrak{B}$ be models of $\Sigma_{n,p}$ of cardinal $\mathfrak{m}$, where $\mathfrak{m}$ is uncountable.

Since $\mathfrak{A}$ and $\mathfrak{B}$ are finite dimensional their fields are both of cardinal $\mathfrak{m}$ and hence, by the result of example 1.5, are isomorphic. But any two vector spaces of the same dimension over isomorphic fields are themselves isomorphic. Thus $\mathfrak{A}$ is isomorphic to $\mathfrak{B}$. This shows that the theory $\Sigma_{n,p}$ is $\mathfrak{m}$-categorical for uncountable $\mathfrak{m}$ and therefore, by Vaught's test, is complete.

It should be noted that, using Kochen's ultralimit theorem, this last result can be extended to arbitrary complete theories of fields. Suppose that Δ is any complete extension of the theory $\{\sigma_F\}$ of fields. Let Δ' be the set of sentences obtained from Δ by *relativizing* the quantifiers to the unary predicate letter F (this notion will be explained in ch. 12). Thus Δ' is the set of sentences which says the same things about the elements satisfying F as Δ says about all elements. Suppose n is an integer and let

$$\Delta_n = \Delta' \cup \{\sigma_{\text{VS}}, \tau_{\text{Dim}(n)}\}.$$

Δ_n is the theory of n-dimensional vector spaces over fields that are models of Δ.

THEOREM 1.18. *The theory Δ_n is complete.*

PROOF. Let

$$\mathfrak{A} = \langle A, F_{\mathfrak{A}}, V_{\mathfrak{A}}, S_{\mathfrak{A}}, P_{\mathfrak{A}}, A_{\mathfrak{A}}, M_{\mathfrak{A}}, 0_{\mathfrak{A}}, \mathbf{0}_{\mathfrak{A}} \rangle$$

and

$$\mathfrak{B} = \langle B, F_{\mathfrak{B}}, V_{\mathfrak{B}}, S_{\mathfrak{B}}, P_{\mathfrak{B}}, A_{\mathfrak{B}}, M_{\mathfrak{B}}, 0_{\mathfrak{B}}, \mathbf{0}_{\mathfrak{B}} \rangle$$

be two models of Δ_n. Then, since Δ is complete

$$\mathfrak{A}' = \langle F_{\mathfrak{A}}, S_{\mathfrak{A}}, P_{\mathfrak{A}}, 0_{\mathfrak{A}} \rangle \text{ and } \mathfrak{B}' = \langle F_{\mathfrak{B}}, S_{\mathfrak{B}}, P_{\mathfrak{B}}, 0_{\mathfrak{B}} \rangle$$

are elementarily equivalent fields. By Kochen's ultralimit theorem there are ultralimits $\mathfrak{A}'_\omega, \mathfrak{B}'_\omega$ of $\mathfrak{A}', \mathfrak{B}'$ respectively, which are isomorphic. Let $\mathfrak{A}_\omega, \mathfrak{B}_\omega$ be the ultralimits of $\mathfrak{A}, \mathfrak{B}$ respectively, with respect to the same ultrafilter sequences. $\mathfrak{A} \prec \mathfrak{A}_\omega$ and $\mathfrak{B} \prec \mathfrak{B}_\omega$ and so $\mathfrak{A}_\omega$ and $\mathfrak{B}_\omega$ are both n-dimensional vector spaces. Since $\mathfrak{A}'_\omega \simeq \mathfrak{B}'_\omega$ they are n-dimensional vector spaces over isomorphic fields and therefore are themselves isomorphic. It follows that $\mathfrak{A} \equiv \mathfrak{B}$.

We have thus shown that all models of Δ_n are elementarily equivalent and so our proof is completed.

Observe however that we cannot apply Vaught's test directly to *infinite* dimensional vector spaces over algebraically closed fields of specified characteristic. For suppose $\mathfrak{A}$ and $\mathfrak{B}$ are two such vector spaces of the same cardinal, say $\aleph_1$. Then it could be that $\mathfrak{A}$ is an $\aleph_0$-dimensional vector space over a field of cardinal $\aleph_1$ while $\mathfrak{B}$ is an $\aleph_1$-dimensional vector space over a field of cardinal $\aleph_0$, and in this case $\mathfrak{A}$ is not isomorphic to $\mathfrak{B}$. However by using the two-cardinal results of ch. 12 we can extend theorem 1.18 to infinite dimensional vector spaces. We show how this can be done after we have developed the necessary machinery (see theorem 12.5.6).

Although Vaught's test is useful, it cannot be applied in all cases. There are theories which are complete but which are not $\mathfrak{m}$-categorical for any infinite $\mathfrak{m}$. We give some examples later. We develop a method which can be used to deal with some of these examples in the next section. We conclude this section with an important fact about completeness which we have already used in the previous chapter.

THEOREM 1.19. LINDENBAUM'S THEOREM

Every theory has a complete consistent extension.

PROOF. Let Σ be a theory. Since Σ is consistent, by the Gödel–Henkin completeness theorem, Σ has a model, say $\mathfrak{A}$. Let $\Delta = Th(\mathfrak{A})$. Then $\Sigma \subseteq \Delta$ and Δ is complete.

Although, as we have just seen, Lindenbaum's theorem is an easy consequence of the Gödel–Henkin completeness theorem, it was originally established before the completeness theorem was proved and was in fact used by Henkin in his proof (HENKIN [1949]) of the completeness theorem. We avoided using it, because we adopted the Boolean algebra method; but see the exercise below.

EXERCISE 1.20. What is the relation between Lindenbaum's theorem and the theorem which states that any set of elements of a Boolean algebra having the fip can be extended to an ultrafilter? Compare our proof of theorem 1.3.4 with Robinson's proof (ROBINSON [1963], p. 88) of Lindenbaum's theorem.

§ 2. Model completeness

In § 1 we considered a theory Σ and the relations that might hold between its various models. The relations we considered were isomorphism (categoricity) and elementary equivalence (completeness). In this section we consider the case when the relation of being an elementary extension holds between those pairs of models of a theory of which one is an extension of the other.

A theory Σ is said to be *model complete* if whenever $\mathfrak{A}$ and $\mathfrak{B}$ are models of Σ and $\mathfrak{A} \subseteq \mathfrak{B}$, then $\mathfrak{A} \prec \mathfrak{B}$. (In future for convenience we will write '$\mathfrak{A} \subseteq \mathfrak{B} \in K$' for '$\mathfrak{A}$ and $\mathfrak{B}$ are in K and $\mathfrak{A} \subseteq \mathfrak{B}$'.)

We first explain the use of the term 'model complete' for this new idea. We have seen that given a language L we can extend it to a language L_γ with constants. Suppose $\mathfrak{A} = \langle A, \{R_\xi : \xi < \alpha\}\rangle$ is a relational structure and A can be well-ordered as a sequence $a = \langle a_\xi : \xi < \gamma\rangle$. Then we denote by $L(\mathfrak{A})$ the language with a sequence of constants $\langle c_\xi : \xi < \gamma\rangle$ that we use to talk about the relational structure $(\mathfrak{A}, a)$.

The *diagram* of $\mathfrak{A}$, which we denote by $D(\mathfrak{A})$, is the set of all *atomic sentences* of $L(\mathfrak{A})$ which hold in $(\mathfrak{A}, a)$ together with the negations of all those atomic sentences that don't hold in $(\mathfrak{A}, a)$. This notion, as well as that of model completeness, is due to Abraham Robinson. The term 'model complete' is now explained by the next theorem.

THEOREM 2.1. *A theory Σ is model complete iff whenever $\mathfrak{A}$ is a model of Σ, $\Sigma \cup D(\mathfrak{A})$ is a complete theory of the language* $L(\mathfrak{A})$.

PROOF. Suppose there is some $\mathfrak{A} \in \boldsymbol{M}(\Sigma)$, such that $\Sigma \cup D(\mathfrak{A})$ is not complete. Then there is some sentence σ of $L(\mathfrak{A})$ such that neither σ nor $\neg\sigma$ is

a consequence of $\Sigma \cup D(\mathfrak{A})$. Suppose, say, σ holds in $(\mathfrak{A},a)$. $\Sigma \cup D(\mathfrak{A}) \cup \{\neg\sigma\}$ is consistent and so has a model; since this model is a model of $D(\mathfrak{A})$, we may suppose this model to be an extension of $(\mathfrak{A},a)$, say $(\mathfrak{B},a)$ with $\mathfrak{A} \subseteq \mathfrak{B}$. Because σ holds in $(\mathfrak{A},a)$ and $\neg\sigma$ holds in $(\mathfrak{B},a)$,

$$(\mathfrak{A},a) \not\equiv (\mathfrak{B},a).$$

Therefore, by theorem 4.1.10, $\mathfrak{B}$ is not an elementary extension of $\mathfrak{A}$. But $\mathfrak{A} \subseteq \mathfrak{B} \in \boldsymbol{M}(\Sigma)$ and so we have shown that Σ is not model complete.

Conversely suppose that whenever $\mathfrak{A} \in \boldsymbol{M}(\Sigma)$, $\Sigma \cup D(\mathfrak{A})$ is complete, and that $\mathfrak{A} \subseteq \mathfrak{B} \in \boldsymbol{M}(\Sigma)$. Since $\Sigma \cup D(\mathfrak{A})$ is complete $(\mathfrak{A},a) \equiv (\mathfrak{B},a)$, and therefore $\mathfrak{A} \prec \mathfrak{B}$. This shows that Σ is model complete.

EXAMPLES

2.2. We give an example of a theory which is complete but not model complete. Let $\mathfrak{A},\mathfrak{B}$ be as in example 4.1.7, and let $\Sigma = Th(\mathfrak{A})$. We show later in this chapter that Σ is complete. Since $\mathfrak{B} \simeq \mathfrak{A}$, $\mathfrak{B}$ is also a model of Σ. Since $\mathfrak{A} \subseteq \mathfrak{B}$ but not $\mathfrak{A} \prec \mathfrak{B}$, Σ is not model complete.

2.3. We give an example of a theory which is model complete but not complete. The theory of an algebraically closed field of unspecified characteristic, Σ_{CF} in the notation of example 5.3.20, is not complete, since for each p, the sentence C_p which says that the field is of characteristic p, can neither be proved nor refuted from Σ_{CF}. However, as we show below, Σ_{CF} is model complete.

We give some more examples of model complete theories later. The last two examples show that completeness does not imply model completeness and, conversely, model completeness does not imply completeness. However they are closely related as we shall soon see.

Let Σ be some theory. A model $\mathfrak{A}$ of Σ is said to be a *prime model* of Σ if every model of Σ has some substructure which is isomorphic to $\mathfrak{A}$.

EXAMPLE 2.4. An ordered set of the same order type as the rationals is a prime model for the theory of densely ordered sets without first or last element.

EXERCISES

2.5. Give an example of a prime model for the theory of a group.

2.6. Give an example of a theory which has no prime model.

We now have

THEOREM 2.7. THE PRIME MODEL TEST

Let Σ be a model complete theory with a prime model $\mathfrak{A}^*$; then Σ is complete.

PROOF. Let $\mathfrak{A},\mathfrak{A}'$ be any two models of Σ. Since $\mathfrak{A}^*$ is a prime model of Σ, $\mathfrak{A},\mathfrak{A}'$ have substructures $\mathfrak{B},\mathfrak{B}'$ respectively, which are isomorphic to $\mathfrak{A}^*$ and hence to each other. But $\mathfrak{B}\subseteq\mathfrak{A}$ and $\mathfrak{A}$ and $\mathfrak{B}$ are models of the model complete theory Σ, and so $\mathfrak{B}\prec\mathfrak{A}$, whence $\mathfrak{A}\equiv\mathfrak{B}$. Similarly $\mathfrak{A}'\equiv\mathfrak{B}'$, and so, because $\mathfrak{B}\simeq\mathfrak{B}'$, $\mathfrak{A}\equiv\mathfrak{A}'$. Thus any two models of Σ are elementarily equivalent and so Σ is complete.

We will give some applications of this theorem when we have shown how to test theories for model completeness. We set about this in the next section.

§ 3. Universal and existential sentences

A formula ϕ is said to be in *prenex normal form*, if it is of the form

$$(Q_1x_1)(Q_2x_2)\ldots(Q_nx_n)\,\psi\,,$$

where each Q_i is a quantifier symbol, either $\exists$ or $\forall$, and ψ is a formula which contains no quantifiers. ψ is called the *matrix* of ϕ. For example, the formula

$$(\forall v_0)(\exists v_2)[P(v_0,v_1)\rightarrow P(v_2,v_0)]$$

is in prenex normal form with matrix $[P(v_0,v_1)\rightarrow P(v_2,v_0)]$.

We state without proof the following well known result:

THEOREM 3.1. *For every formula ϕ there is a formula ψ containing the same predicate letters and the same free variables as ϕ, which is in prenex normal form and is such that $\vdash\phi\leftrightarrow\psi$.*

A proof of this theorem can be found, for example, in MENDELSON [1964] p. 87.

By an *open formula* we mean a formula which contains no quantifiers. Such formulas are sometimes called *quantifier free formulas*. We let $\forall$ be the set of all formulas of the form

$$(\forall x_1)\ldots(\forall x_n)\,\psi\,,$$

where ψ is an open formula. Thus $\forall$ is the set of all prenex normal form formulas containing only universal quantifiers. The formulas in $\forall$ are called *universal formulas*. Similarly $\exists$ is the set of all formulas of the form

$$(\exists x_1)\ldots(\exists x_n)\,\psi\,,$$

where ψ is an open formula. The formulas in $\exists$ are called *existential formulas*.

Notice that, provided the x_i and y_j are distinct,

$$\vdash [(\forall x_1)\ldots(\forall x_m)\,\phi \wedge (\forall y_1)\ldots(\forall y_n)\,\psi] \leftrightarrow (\forall x_1)\ldots(\forall x_m)(\forall y_1)\ldots(\forall y_n)[\phi \wedge \psi]$$

and thus given any finite number of universal formulas $\phi_1,\ldots,\phi_n$, their conjunction $\phi_1 \wedge \cdots \wedge \phi_n$ is equivalent to a universal formula. Similarly a conjunction of existential formulas is equivalent to an existential formula.

EXERCISE 3.2. If ϕ is equivalent to a universal formula, show that $\neg\phi$ is equivalent to an existential formula, and vice versa.

We are now going to generalize the notions of elementary equivalence and embeddability by restricting our attention to universal and existential formulas only.

We say that $\mathfrak{B}$ *covers* $\mathfrak{A}$ *universally* and we write $\mathfrak{A} \forall \mathfrak{B}$, if for all sentences $\sigma \in \forall$, $\mathfrak{A} \models \sigma$ implies $\mathfrak{B} \models \sigma$. Thus $\mathfrak{A} \forall \mathfrak{B}$ iff $Th(\mathfrak{A}) \cap \forall \subseteq Th(\mathfrak{B}) \cap \forall$.

We say that $\mathfrak{B}$ *covers* $\mathfrak{A}$ *existentially* and we write $\mathfrak{A} \exists \mathfrak{B}$, if for all sentences $\sigma \in \exists$, $\mathfrak{A} \models \sigma$ implies $\mathfrak{B} \models \sigma$.

$\mathfrak{A}$ is said to be a *universal substructure* of $\mathfrak{B}$ if $\mathfrak{A} \subseteq \mathfrak{B}$ and for any universal formula ϕ and any $x \in A^\omega$,

$$\mathfrak{A} \models_x \phi \quad \text{implies} \quad \mathfrak{B} \models_x \phi .$$

If $\mathfrak{A}$ is a universal substructure of $\mathfrak{B}$ we write $\mathfrak{A} \subseteq_\forall \mathfrak{B}$.

$\mathfrak{A}$ is said to be an *existential substructure* of $\mathfrak{B}$, and we write $\mathfrak{A} \subseteq_\exists \mathfrak{B}$, if $\mathfrak{A} \subseteq \mathfrak{B}$ and for any existential formula ϕ and any $x \in A^\omega$,

$$\mathfrak{A} \models_x \phi \quad \text{implies} \quad \mathfrak{B} \models_x \phi .$$

As we see below, if $\mathfrak{A}$ is a substructure of $\mathfrak{B}$ then $\mathfrak{A}$ is automatically an existential substructure of $\mathfrak{B}$.

The following facts are almost immediate consequences of the definitions.

LEMMA 3.3. (a). $\mathfrak{A} \forall \mathfrak{B}$ *iff* $\mathfrak{B} \exists \mathfrak{A}$.

(b). $\mathfrak{A} \subseteq_\forall \mathfrak{B}$ *iff* $\mathfrak{A} \subseteq \mathfrak{B}$ *and for any* $\phi \in \exists$ *and any* $x \in A^\omega$, $\mathfrak{B} \models_x \phi$ *implies* $\mathfrak{A} \models_x \phi$.

PROOF. From the definitions and exercise 3.2.

EXERCISES

3.4. Show that if $\mathfrak{A} \subseteq \mathfrak{B}$, ϕ is an open formula and $x \in A^\omega$, then $\mathfrak{A} \models_x \phi$ implies $\mathfrak{B} \models_x \phi$.

3.5. Show that $\mathfrak{A} \prec \mathfrak{B}$ implies $\mathfrak{A} \subseteq_\forall \mathfrak{B}$ and $\mathfrak{A} \subseteq_\exists \mathfrak{B}$.

3.6. Show that $\forall$, $\exists$, $\subseteq_\forall$ and $\subseteq_\exists$ are all transitive relations.

LEMMA 3.7. *If* $\mathfrak{A} \subseteq \mathfrak{B}$ *then* $\mathfrak{A} \subseteq_\exists \mathfrak{B}$.

PROOF. Suppose $\mathfrak{A} \subseteq \mathfrak{B}$, ϕ is an existential formula and for some $x \in A^\omega$, $\mathfrak{A} \models_x \phi$. ϕ is of the form

$$(\exists v_0) \ldots (\exists v_n)\, \psi$$

where ψ is an open formula. Since $\mathfrak{A} \models_x \phi$, $\mathfrak{A} \models_{x'} \psi$, where x' is some sequence in A^ω differing from x in at most the first $n+1$ places. By exercise 3.4, $\mathfrak{B} \models_{x'} \psi$, and so $\mathfrak{B} \models_x \phi$. This shows that $\mathfrak{A} \subseteq_\exists \mathfrak{B}$.

A sentence σ is said to be *persistent with respect to extensions* if whenever $\mathfrak{A} \models \sigma$ and $\mathfrak{A} \subseteq \mathfrak{B}$, then $\mathfrak{B} \models \sigma$. The previous lemma shows that existential sentences are persistent with respect to extensions.

The next two lemmas are exact analogues of lemmas 8.1.1 and 8.1.3 of the previous chapter.

LEMMA 3.8. $\mathfrak{A} \forall \mathfrak{B}$ *iff there is an ultrafilter pair* $\langle I,F \rangle$ *such that* $\mathfrak{B}$ *can be embedded in* $\mathfrak{A}^I/F$.

LEMMA 3.9. *The map* $g: \mathfrak{A} \to \mathfrak{B}$ *is an embedding such that* $g(\mathfrak{A}) \subseteq_\forall \mathfrak{B}$ *iff there is an ultrafilter pair* $\langle I,F \rangle$ *and an embedding* $h: \mathfrak{B} \to \mathfrak{A}^I/F$ *such that diagram* 3.a *is commutative, where d is the canonical embedding.*

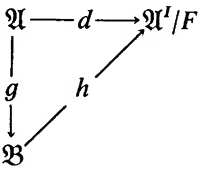

Diagram 3.a

EXERCISE 3.10. Write out proofs of lemmas 3.8 and 3.9. (Follow the proofs of lemmas 8.1.1 and 8.1.3, using lemma 3.3(a) and (b) respectively, instead of the symmetry of the relation $\equiv$ at the obvious places.)

Lemma 3.9 has the following consequences.

COROLLARY 3.11. *If* $\mathfrak{A} \subseteq \mathfrak{B}$, $\mathfrak{A} \subseteq_\forall \mathfrak{B}$ *iff there is an ultrafilter pair* $\langle I,F \rangle$ *and an embedding* $h: \mathfrak{B} \to \mathfrak{A}^I/F$ *such that* $h \restriction \mathfrak{A}$, *the restriction of h to* $\mathfrak{A}$, *is the canonical embedding d.*

PROOF. This is just the special case of lemma 3.9 when g is the injection map.

COROLLARY 3.12. $\mathfrak{A} \subseteq_{\forall} \mathfrak{B}$ *iff there is a relational structure* $\mathfrak{C}$ *such that* $\mathfrak{A} \subseteq \mathfrak{B} \subseteq \mathfrak{C}$ *and* $\mathfrak{A} \prec \mathfrak{C}$.

PROOF. Suppose $\mathfrak{A} \subseteq_{\forall} \mathfrak{B}$, then by corollary 3.11 there is an ultrafilter pair $\langle I,F \rangle$ such that $\mathfrak{B}$ can be embedded in $\mathfrak{A}^I/F$. Hence some extension $\mathfrak{C}$ of $\mathfrak{B}$ is isomorphic to $\mathfrak{A}^I/F$. But then $\mathfrak{A} \subseteq \mathfrak{B} \subseteq \mathfrak{C}$ and since $\mathfrak{A} \prec \mathfrak{A}^I/F$, also $\mathfrak{A} \prec \mathfrak{C}$.

Conversely suppose that such a structure $\mathfrak{C}$ exists. Since $\mathfrak{B} \subseteq \mathfrak{C}$, $\mathfrak{B} \subseteq_{\exists} \mathfrak{C}$. Now suppose $\phi \in \forall$ and $\mathfrak{A} \models_x \phi$. Since $\mathfrak{A} \prec \mathfrak{C}$, $\mathfrak{C} \models_x \phi$, and since $\mathfrak{B} \subseteq_{\exists} \mathfrak{C}$, by lemma 3.3, $\mathfrak{B} \models_x \phi$. Therefore $\mathfrak{A} \subseteq_{\forall} \mathfrak{B}$.

In the next section we are going to use these results to obtain a test for model completeness in terms of existential formulas.

§4. A test for model completeness

The main result of this section stems from the following theorem:

THEOREM 4.1. *Let* K *be a class of relational structures closed under isomorphism and the formation of ultrapowers, then for all* $\mathfrak{A} \subseteq \mathfrak{B} \in K$, $\mathfrak{A} \prec \mathfrak{B}$ *iff*

for all $\mathfrak{A} \subseteq \mathfrak{B} \in K$, *there is an ultrafilter pair* $\langle I,F \rangle$ *and an embedding* $h: \mathfrak{B} \to \mathfrak{A}^I/F$ *such that* $h \restriction \mathfrak{A} = d$, *the canonical embedding of* $\mathfrak{A}$ *into* $\mathfrak{A}^I/F$. [Θ]

PROOF. The necessity of the condition [Θ] follows from corollary 3.11.

Conversely suppose that [Θ] holds. Since K is closed under isomorphism the following condition also holds:

Whenever $\mathfrak{A}, \mathfrak{B} \in K$ *and* $g: \mathfrak{A} \to \mathfrak{B}$ *is an embedding there is an ultrafilter pair* $\langle I,F \rangle$ *and an embedding* $h: \mathfrak{B} \to \mathfrak{A}^I/F$ *such that diagram* 4.a *is commutative, where* d *is the canonical embedding:* [Δ]

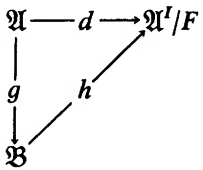

Diagram 4.a

Now suppose $\mathfrak{A} \subseteq \mathfrak{B} \in K$; we will show that $\mathfrak{A} \prec \mathfrak{B}$. Let $g_0: \mathfrak{A} \to \mathfrak{B}$ be the injection map. Then by [Δ] there is an ultrafilter pair $\langle I_1, F_1 \rangle$ and an embedding $h_1: \mathfrak{B} \to \mathfrak{A}^{I_1}/F_1 = \mathfrak{A}_1$, say, such that diagram 4.b is commutative, where d_1 is the canonical embedding.

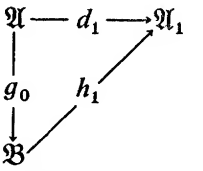

Diagram 4.b

Since K is closed under the formation of ultrapowers, $\mathfrak{A}_1 \in K$, and so, by [Δ], since $h_1 : \mathfrak{B} \to \mathfrak{A}_1$ is an embedding, there is an ultrafilter pair $\langle J_1, G_1 \rangle$ and an embedding $g_1 : \mathfrak{A}_1 \to \mathfrak{B}^{J_1}/G_1 = \mathfrak{B}_1$, say, such that diagram 4.c is commutative, where d_1, e_1 are the canonical embeddings,

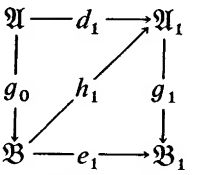

Diagram 4.c

Clearly, just as in the proof of Kochen's ultralimit theorem, this diagram can be extended to an infinite commutative diagram. Let $\mathfrak{A}_\omega, \mathfrak{B}_\omega$ be the ultralimits of $\mathfrak{A}, \mathfrak{B}$ respectively, with respect to the ultrafilter sequences $\langle \langle I_n, F_n \rangle : n \in \omega \rangle, \langle \langle J_n, G_n \rangle : n \in \omega \rangle$ so obtained. Again, as in the proof of the ultralimit theorem, we can define the map $g_\omega : \mathfrak{A}_\omega \to \mathfrak{B}_\omega$ such that g_ω is an isomorphism and diagram 4.d is commutative, where d and e are the projections,

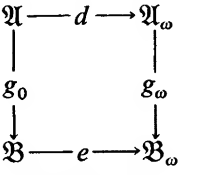

Diagram 4.d

That is, $g_\omega \cdot d = e \cdot g_0$. Thus if $x \in A^\omega$, and ϕ is a formula then

$$\begin{aligned}
\mathfrak{A} \vDash_x \phi \quad &\text{iff} \quad \mathfrak{A}_\omega \vDash_{d(x)} \phi \\
&\text{iff} \quad \mathfrak{B}_\omega \vDash_{g_\omega(d(x))} \phi \\
&\text{iff} \quad \mathfrak{B}_\omega \vDash_{e(g_0(x))} \phi \\
&\text{iff} \quad \mathfrak{B} \vDash_{g_0(x)} \phi \\
&\text{iff} \quad \mathfrak{B} \vDash_x \phi, \quad \text{since } g_0 \text{ is the injection map.}
\end{aligned}$$

This completes the proof that $\mathfrak{A} \prec \mathfrak{B}$ and hence of the theorem.

Now suppose $K = \boldsymbol{M}(\Sigma)$, that is that K is the class of all models of some theory Σ. Then K is certainly closed under isomorphism and also under the formation of ultrapowers. But the condition that whenever $\mathfrak{A} \subseteq \mathfrak{B} \in K$, $\mathfrak{A} \prec \mathfrak{B}$ is just the condition that Σ is model complete. Thus we obtain the following corollaries of the theorem we have just proved.

COROLLARY 4.2. *A theory Σ is model complete iff whenever $\mathfrak{A} \subseteq \mathfrak{B} \in \boldsymbol{M}(\Sigma)$ there is an ultrafilter pair $\langle I,F \rangle$ and an embedding $h: \mathfrak{B} \to \mathfrak{A}^I/F$ such that $h \restriction \mathfrak{A} = d$, the canonical embedding.*

COROLLARY 4.3. *Let K be a class of relational structures closed under isomorphism and the formation of ultrapowers. Then, for all $\mathfrak{A} \subseteq \mathfrak{B} \in K$, $\mathfrak{A} \prec \mathfrak{B}$ iff for all $\mathfrak{A} \subseteq \mathfrak{B} \in K$, $\mathfrak{A} \subseteq_\forall \mathfrak{B}$.*

PROOF. This is an immediate consequence of corollary 3.11 and theorem 4.1.

COROLLARY 4.4. *A theory Σ is model complete iff for all models $\mathfrak{A}, \mathfrak{B}$ of Σ such that $\mathfrak{A} \subseteq \mathfrak{B}$, $\mathfrak{A} \subseteq_\forall \mathfrak{B}$.*

Thus we have shown that the condition that Σ be model complete can be reduced from the condition

$$\mathfrak{A} \subseteq \mathfrak{B} \in \boldsymbol{M}(\Sigma) \quad \text{implies} \quad \mathfrak{A} \prec \mathfrak{B}$$

to the apparently weaker condition

$$\mathfrak{A} \subseteq \mathfrak{B} \in \boldsymbol{M}(\Sigma) \quad \text{implies} \quad \mathfrak{A} \subseteq_\forall \mathfrak{B}.$$

It follows that in order to test for model completeness we need only look at universal sentences. We are now going to show that the condition can be made even more simple.

A formula ϕ is said to be *primitive* if ϕ is an existential formula whose matrix is a conjunction of atomic formulas and negations of atomic formulas. We shall obtain a test for model completeness involving primitive formulas only.

THEOREM 4.5. *Let K be a class of relational structures closed under isomorphism and the formation of ultrapowers. Then for all $\mathfrak{A} \subseteq \mathfrak{B} \in K$, $\mathfrak{A} \prec \mathfrak{B}$ iff for all $\mathfrak{A} \subseteq \mathfrak{B} \in K$, for all primitive formulas ϕ and all sequences $x \in A^\omega$, if $\mathfrak{B} \models_x \phi$ then $\mathfrak{A} \models_x \phi$.*

PROOF. By lemma 3.3(b), $\mathfrak{A} \subseteq_\forall \mathfrak{B}$ iff $\mathfrak{A} \subseteq \mathfrak{B}$ and for all $\phi \in \exists$, $x \in A^\omega$, if $\mathfrak{B} \models_x \phi$ then $\mathfrak{A} \models_x \phi$. Therefore, by corollary 4.3, it is sufficient to prove that the following two statements are equivalent:

(A). *For each formula* $\phi \in \exists$, *and each* $x \in A^\omega$, *if* $\mathfrak{B} \models_x \phi$ *then* $\mathfrak{A} \models_x \phi$.

(B). *For each primitive formula* ϕ, *and each* $x \in A^\omega$, *if* $\mathfrak{B} \models_x \phi$ *then* $\mathfrak{A} \models_x \phi$.

Clearly (A) implies (B).

Now suppose (B) holds. Let $\phi \in \exists$, $x \in A^\omega$ and suppose $\mathfrak{B} \models_x \phi$; we show that $\mathfrak{A} \models_x \phi$. For some open formula χ, ϕ is

$$(\exists v_{i_1}) \ldots (\exists v_{i_n})\, \chi .$$

Let χ be written in disjunctive normal form (see e.g. MENDELSON [1964] p. 27) as $\chi_1 \vee \cdots \vee \chi_r$. Let ψ_i be

$$(\exists v_{i_1}) \ldots (\exists v_{i_n})\, \chi_i \quad \text{for} \quad 1 \leqslant i \leqslant r .$$

Each ψ_i is a primitive formula and

$$\vdash \phi \leftrightarrow (\psi_1 \vee \cdots \vee \psi_r) .$$

Since $\mathfrak{B} \models_x \phi$,

$$\mathfrak{B} \models_x \psi_1 \vee \cdots \vee \psi_r$$

and so for some i, $1 \leqslant i \leqslant r$,

$$\mathfrak{B} \models_x \psi_i .$$

Hence by our assumption (B), $\mathfrak{A} \models_x \psi_i$ and so $\mathfrak{A} \models_x \phi$. This proves the theorem.

COROLLARY 4.6. ROBINSON'S MODEL COMPLETENESS TEST
A theory Σ *is model complete iff whenever* $\mathfrak{A} \subseteq \mathfrak{B} \in \boldsymbol{M}(\Sigma)$ *then for any primitive formula* ϕ *and any* $x \in A^\omega$, *if* $\mathfrak{B} \models_x \phi$ *then* $\mathfrak{A} \models_x \phi$.

The advantage of this last result is obvious. It makes testing for model completeness a very much simpler matter, as we have only to look at primitive formulas. We devote the next section to some applications of this result.

§ 5. Applications of Robinson's test

In this section we give some applications of corollary 4.6, Robinson's model completeness test, to some specific algebraic theories.

(a). *Densely ordered sets without first or last element*

In this case the appropriate language L has one binary predicate letter P to represent the order relation. Let σ_{DOS} be the sentence of example 5.3.10 which characterizes densely ordered sets and let

$$\Sigma = \{\sigma_{\mathrm{DOS}}, (\forall v_0)[(\exists v_1)\, P(v_0, v_1) \wedge (\exists v_1)\, P(v_1, v_0)]\} .$$

Σ is the theory of densely ordered sets without first or last element. We have already proved that this theory is complete; we now show that it is model complete.

Suppose then that $\mathfrak{A},\mathfrak{B}$ are models of Σ, $\mathfrak{A}\subseteq\mathfrak{B}$, $a\in A^{\omega}$ and ϕ is a primitive formula such that $\mathfrak{B}\models_a\phi$. By Robinson's test it will be sufficient to show that $\mathfrak{A}\models_a\phi$. Let ψ be the matrix of ϕ. Let $v_{i_0},\ldots,v_{i_s}$ be the free variables occurring in ϕ and let $v_{j_0},\ldots,v_{j_t}$ be the bound variables occurring in ϕ. By hypothesis there exist $b_{j_0},\ldots,b_{j_t}$ in B such that

$$\mathfrak{B}\models\psi[a_{i_0},\ldots,a_{i_s},b_{j_0},\ldots,b_{j_t}].$$

Our aim is to show that there exist $a_{j_0},\ldots,a_{j_s}$ in A such that

$$\mathfrak{A}\models\psi[a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}]$$

If $n=i_k$ let $c_n=a_{i_k}$ and if $n=j_l$ let $c_n=b_{j_l}$. Also let $<$ be the value of P in $\mathfrak{B}$. The matrix ψ of ϕ is a conjunction of formulas of the forms

$$P(v_m,v_n),\quad \neg P(v_m,v_n),\quad v_m=v_n,\quad \neg v_m=v_n.$$

We eliminate all the formulas of the last three kinds by carrying out the following steps in turn:

(i). If $\neg v_m=v_n$ occurs, $c_m\neq c_n$ and hence either $c_m<c_n$ or $c_n<c_m$. If $c_m<c_n$, we replace $\neg v_m=v_n$ by $P(v_m,v_n)$ and otherwise we replace it by $P(v_n,v_m)$.

(ii). If $\neg P(v_m,v_n)$ occurs, $c_m\not<c_n$ and hence either $c_n<c_m$ or $c_m=c_n$. If $c_n<c_m$ we replace $\neg P(v_m,v_n)$ by $P(v_n,v_m)$ and otherwise by $v_m=v_n$.

(iii). If $v_m=v_n$ occurs, we replace all other occurrences of v_n in ψ by v_m and we drop the formula $v_m=v_n$.

These steps will lead to a new open formula ψ'. ψ' will be a conjunction of atomic formulas of the form

$$P(v_m,v_n).$$

It can be seen from the method we have used to construct ψ' that

$$\mathfrak{B}\models\psi'[a_{i_0},\ldots,a_{i_s},b_{j_0},\ldots,b_{j_t}],$$

and that if

$$\mathfrak{A}\models\psi'[a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}],$$

then also

$$\mathfrak{A}\models\psi[a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}].$$

Clearly, without loss of generality, we can assume that

$$b_{j_0}<\cdots<b_{j_t}.$$

Suppose that in $\mathfrak{B}$ the set $\{a_{i_0},\ldots,a_{i_s},b_{j_0},\ldots,b_{j_t}\}$ is ordered as, say,

$$b_{j_0}<b_{j_1}<a_{i_0}<\cdots<b_{j_k}<a_{i_l}<b_{j_{k+1}}<\cdots<a_{i_s}<b_{j_t} \tag{1}$$

Since $\mathfrak{A}$ is densely ordered, there exist $a_{j_0},\ldots,a_{j_t}$ in A such that

$$a_{j_0}<a_{j_1}<a_{i_0}<\cdots<a_{j_k}<a_{i_l}<a_{j_{k+1}}<\cdots<a_{i_s}<a_{j_t} \tag{2}$$

ψ' is a conjunction of atomic formulas of the form $P(v_m,v_n)$ and the elements satisfying conditions (1) satisfy ψ' in $\mathfrak{B}$. Hence the elements satisfying (2) also satisfy ψ' in $\mathfrak{A}$. That is

$$\mathfrak{A}\models\psi'[a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}].$$

Therefore by the remarks above

$$\mathfrak{A}\models_a\phi.$$

We have therefore proved that Σ is model complete.

Let η be the order type of the rationals. Any set of order type η is a prime model of Σ. Therefore we can conclude from the prime model test that Σ is complete, a result that we deduced earlier from Vaught's test.

(b). *Densely ordered sets with first but no last element*

Let

$$\Sigma_1=\{\sigma_{\mathrm{DOS}},(\exists v_0)(\forall v_1)[\neg P(v_1,v_0)],(\forall v_0)(\exists v_1)\,P(v_0,v_1)\}.$$

Σ_1 is the theory of densely ordered sets with first but no last element. Σ_1 is *not* model complete. The simple proof of this left as an exercise. We can make Σ_1 model complete by adding a new unary predicate letter F to the language L to form the extended language L'. Let

$$\Sigma_1'=\Sigma_1\cup\{(\forall v_0)[F(v_0)\leftrightarrow\neg(\exists v_1)\,P(v_1,v_0)]\}.$$

Thus $F(x)$ 'says' that x is the first element. We show that Σ_1' is model complete.

Again suppose $\mathfrak{A},\mathfrak{B}$ are models of Σ_1', $\mathfrak{A}\subseteq\mathfrak{B}$, ϕ is a primitive formula with matrix ψ, $a\in A^\omega$ and $\mathfrak{B}\models_a\phi$.

This time the matrix ψ of ϕ will be a conjunction of formulas of the forms

$$P(v_m,v_n),\quad\neg P(v_m,v_n),\quad v_m=v_n,\quad\neg v_m=v_n,\quad F(v_m),\quad\neg F(v_m).$$

$\mathfrak{A}$ has a first element. Without loss of generality we can assume this first element is a_0 and that the variable v_0 does not occur in ϕ. $\mathfrak{A}\models F[a_0]$ and $\mathfrak{A}\subseteq\mathfrak{B}$, therefore, because $F(v_0)$ is an atomic formula of L', a_0 is also the

first element of $\mathfrak{B}$. Note that we could not draw this conclusion if we were working with the original language L.

We now replace any occurrence of $F(v_m)$ in ψ by $v_0 = v_m$, and any occurrence of $\neg F(v_m)$ by $\neg v_0 = v_m$ and then we carry out the same replacements as in case (a). It can be seen that the argument used in (a) works here as well and therefore we can conclude that Σ_1' is model complete.

Any set of order type $1+\eta$ is a prime model for Σ_1'. Therefore we can conclude that Σ_1' is complete. From this we can deduce that Σ_1 is complete as follows:

Let $\sigma_1,\ldots,\sigma_n$ be a proof of σ_n from the assumptions Σ_1' and suppose σ_n is a sentence of L, so that σ_n does not contain F. For $i \leqslant n$, let σ_i^* be the result of replacing all occurrences of $F(v_m)$ in σ_i by $\neg(\exists v_1) P(v_1, v_m)$. Since under this transformation the sentence

$$(\forall v_0)[F(v_0) \leftrightarrow \neg(\exists v_1) P(v_1, v_0)],$$

becomes the sentence

$$(\forall v_0)[\neg(\exists v_1) P(v_1, v_0) \leftrightarrow \neg(\exists v_1) P(v_1, v_0)],$$

which is a provable sentence of L, it can easily be seen that all the sentences σ_i^* are provable from the assumptions Σ_1. In particular $\sigma_n^* = \sigma_n$ and so $\Sigma_1 \vdash \sigma_n$.

Thus, for sentences σ_n of L, $\Sigma_1' \vdash \sigma_n$ implies $\Sigma_1 \vdash \sigma_n$. Since Σ_1' is complete it therefore follows that Σ_1 is also complete.

(c). *Densely ordered sets with last but not first element*

(d). *Densely ordered sets with first and last element*

These two cases can be dealt with just like (b) and in each case we can conclude that theories of such ordered sets are complete.

(e). *Discretely ordered sets with no first or last element*

Again the appropriate language L has one binary predicate letter P. Let σ_O be the sentence (described in example 5.3.6) which says that P is a total ordering. Let σ_{OD} be the conjunction of σ_O and the two sentences

$$(\forall v_0)\{(\exists v_1) P(v_1, v_0) \rightarrow (\exists v_1)[P(v_1, v_0) \wedge (\forall v_2) \neg [P(v_1, v_2) \wedge P(v_2, v_0)]]\},$$

'every element with predecessors has an immediate predecessor',

$$(\forall v_0)\{(\exists v_1) P(v_0, v_1) \rightarrow (\exists v_1)[P(v_0, v_1) \wedge (\forall v_2) \neg [P(v_0, v_2) \wedge P(v_2, v_1)]]\},$$

'every element with successors has an immediate successor'.

Let

$$\Sigma_2 = \{\sigma_{\mathrm{OD}}, (\forall v_0)[(\exists v_1)\, P(v_0,v_1) \wedge (\exists v_1)\, P(v_1,v_0)]\}.$$

Σ_2 is the theory of discretely ordered sets with no first or last element. One example of a model of Σ_2 is the integers with the natural ordering:

$$\cdots < -3 < -2 < -1 < 0 < +1 < +2 < +3 < \cdots.$$

This linearly ordered set is of order type $\omega^* + \omega$ and can easily be seen to be a prime model for Σ_2, let it be $\mathfrak{A}_0$.

Another model is given by the integers and the number $+\frac{1}{2}$ with their natural ordering:

$$\cdots < -3 < -2 < -1 < 0 < +\tfrac{1}{2} < +1 < +2 < +3 < \cdots.$$

This model, say $\mathfrak{A}_1$, is an extension of $\mathfrak{A}_0$ but since 1 is the immediate successor of 0 in $\mathfrak{A}_0$ but not in $\mathfrak{A}_1$, $\mathfrak{A}_1$ is not an elementary extension of $\mathfrak{A}_0$, thus Σ_2 is *not* model complete.

Once again we can get over this difficulty by adding a new predicate letter to L. This time we add a new binary predicate letter S and we let

$$\Sigma_2' = \Sigma_2 \cup \{(\forall v_0)(\forall v_1)[S(v_0,v_1) \leftrightarrow \neg(\exists v_2)[P(v_0,v_2) \wedge P(v_2,v_1)]]\}.$$

Thus $S(x,y)$ says that 'if $x<y$ then y is the immediate successor of x'. We will show that Σ_2' is model complete.

Let $\mathfrak{A},\mathfrak{B}$ be models of Σ_2' with $\mathfrak{A} \subseteq \mathfrak{B}$. Suppose ϕ is a primitive formula with matrix ψ, $a \in A^\omega$ and $\mathfrak{B} \vDash_a \phi$. Once again it will be sufficient to prove that $\mathfrak{A} \vDash_a \phi$. Clearly,

$$\Sigma_2' \vdash (\exists v_0)\ldots(\exists v_n)[\sigma \wedge \neg S(v_i,v_j)] \leftrightarrow (\exists v_0)\ldots(\exists v_n)(\exists v_{n+1})[\sigma \wedge P(v_i,v_{n+1}) \wedge P(v_{n+1},v_j)],$$

and therefore we can assume that there is no formula of the form $\neg S(v_m,v_n)$ in ψ, and thus that ψ is a conjunction of formulas of the forms

$$S(v_m,v_n),\quad P(v_m,v_n),\quad \neg P(v_m,v_n),\quad v_m = v_n,\quad \neg v_m = v_n.$$

Using exactly the same method as in (a) we can eliminate all formulas of the last three forms and hence arrive at an open formula ψ', which is a conjunction of formulas of the forms

$$S(v_m,v_n),\quad P(v_m,v_n),$$

such that, using the same notation as in (a),

$$\mathfrak{B} \vDash \psi'[a_{i_0},\ldots,a_{i_s},b_{j_0},\ldots,b_{j_t}],$$

and if

$$\mathfrak{A} \models \psi' [a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}],$$

then also

$$\mathfrak{A} \models \psi [a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}].$$

Suppose that in $\mathfrak{B}$,

$$b_{j_0} < b_{j_1} < a_{i_0} < \cdots < b_{j_k} < a_{i_l} < b_{j_{k+1}} < \cdots < a_{i_s} < b_{j_t}. \tag{1}$$

It will be sufficient to show that we can find $a_{j_0},\ldots,a_{j_t}$ in A so that the same order relations hold when we replace each b_j by the corresponding a_j and, since ψ' also contains formulas of the form $S(v_m,v_n)$, also that whenever one element of (1) is an immediate successor of another, the same relation holds between the corresponding elements of the new sequence.

An m-tuple $\langle x_1,\ldots,x_m\rangle$ will be called a *chain* if whenever $2 \leqslant i \leqslant m$, x_i is the immediate successor of x_{i-1}. Since S is a predicate letter of L′ and $\mathfrak{A} \subseteq \mathfrak{B}$, if $\langle x_1,\ldots,x_m\rangle$ is a chain in $\mathfrak{A}$, then it is also a chain in $\mathfrak{B}$. Also if $\langle x_1,\ldots,x_m\rangle$ is a chain in $\mathfrak{B}$ one of whose elements is in A, then it is easily seen that all the elements of the chain are in A.

Since A has no first or last elements it is easy to replace those b_j in (i) that occur either before all the a_i or after them all by an a_j so as to meet the desired conditions. So we need only consider those b_j lying between a_i's. Say, for example,

$$a_{i_m} < b_{j_n} < \cdots < b_{j_{n+p}} < a_{i_{m+1}}$$

If a_{i_m}, $a_{i_{m+1}}$ are end points of a chain in $\mathfrak{A}$ then by the remarks above all the elements between them in B are in A. Thus each $b_{j_{n+l}}$, $0 \leqslant l \leqslant p$, is in A and so we can let $a_{j_{n+l}} = b_{j_{n+l}}$.

If $a_{i_m}, a_{i_{m+1}}$ are not end points of a chain in $\mathfrak{A}$ it must be that there are infinitely many elements in A between a_{i_m} and $a_{i_{m+1}}$. But then there is no problem in finding $a_{j_n},\ldots,a_{j_{n+p}}$ between a_{i_m} and $a_{i_{m+1}}$ to meet the desired conditions.

Thus in any case we can find $a_{j_0},\ldots,a_{j_t}$ in A such that

$$\mathfrak{A} \models \psi' [a_{i_0},\ldots,a_{i_s},a_{j_0},\ldots,a_{j_t}]$$

and therefore

$$\mathfrak{A} \models_a \phi.$$

We can therefore conclude that Σ_2' is model complete. Since Σ_2' has a prime model it follows that Σ_2' is complete. As in (b) we can draw the conclusion that Σ_2 is complete.

(f). *Discretely ordered sets with first but no last element*

If $\Sigma_3 = \{\sigma_{OD}, (\exists v_0)(\forall v_1)\neg P(v_1,v_0), (\forall v_0)(\exists v_1)P(v_0,v_1)\}$ then Σ_3 is the theory of discretely ordered sets with first but no last element. We have already seen, in example 4.1.7, that this theory is not model complete. However by using the same device as in (e) and the device used in (b) for coping with first elements we can find an extension Σ_3' of Σ_3 in an extended language which is model complete. Since any model of order type ω is a prime model of Σ_3' we can conclude that Σ_3' is complete. It follows as before that Σ_3 is complete. We can conclude from this that ordered sets of order types

$$\omega \quad \text{and} \quad \omega + \omega^* + \omega$$

are elementarily equivalent. This is the result promised in example 4.1.4.

$M(\Sigma_3)$, the class of all models of Σ_3, is the class of all discretely ordered sets with first but no last element. It is easily seen that ordered sets of order type ω are the only well-ordered sets in $M(\Sigma_3)$. Since Σ_3 is complete, $M(\Sigma_3)$ is closed under elementary equivalence but contains sets which are not well-ordered. This gives another proof of lemma 7.1.13.

(g). *Discretely ordered sets with last but no first element*

This case is dealt with like (f) and so the theory of such sets is also complete.

(h). *Infinite discretely ordered sets with first and last elements*

The theory

$$\Sigma_4 = \{\sigma_{OD}, (\exists v_1)(\forall v_0)\neg P(v_1,v_0), (\exists v_1)(\forall v_0)\neg P(v_0,v_1)\}$$

of discretely ordered sets with first and last elements is not complete since it has models of all finite cardinals >0 and also infinite models. Thus none of the sentences $\exists!^n$ or their negations is deducible from Σ_4.

However, by the methods used in (f) it can be shown that the theory

$$\Sigma_5 = \Sigma_4 \cup \{\exists^{\geq n} : n \in \omega\}$$

of infinite discretely ordered sets with first and last elements is complete.

(j). *Algebraically closed fields*

Let Σ_{CF} be the theory of algebraically closed fields described in example 5.3.20; we shall show that Σ_{CF} is model complete.

Suppose $\mathfrak{A}, \mathfrak{B}$ are models of Σ_{CF}, $\mathfrak{A} \subseteq \mathfrak{B}$, ϕ is a primitive formula, $a \in A^\omega$ and $\mathfrak{B} \models_a \phi$.

'$\mathfrak{B} \models_a \phi$' says that there exists in B a solution of a finite set of conditions

of the following sorts

$$u=v,\ u\neq v,\ u+v=w,\ u+v\neq w,\ u\cdot v=w,\ u\cdot v\neq w \qquad (*)$$

where the letters u,v,w denote either elements of A (these will be the elements a_n of the sequence a, corresponding to the free variables of ϕ) or unknowns (corresponding to the bound variables of ϕ). Let the elements of A be $\{a_i: 1\leqslant i\leqslant k\}$ and the unknowns be $\{y_i: 1\leqslant i\leqslant m\}$. The assumption that $\mathfrak{B}\models_a\phi$ is equivalent to the assumption that we can find solutions

$$y_i=b_i,\quad 1\leqslant i\leqslant m$$

in B to satisfy the conditions $(*)$. Our aim is to find solutions in A of these conditions.

The proof that we can find such a solution will involve some knowledge of the theory of fields which can be obtained from, for example, VAN DER WAERDEN [1953].

We first show that we can find solutions in A in the special case when $\mathfrak{B}$ is of degree of transcendence one over $\mathfrak{A}$.

Let D be the diagram of $\mathfrak{A}$. D is a set of sentences of the language $\mathrm{L}(\mathfrak{A})$. Let L' be the language obtained from $\mathrm{L}(\mathfrak{A})$ by adding a new constant symbol c.

For each polynomial

$$p(x)=a_0+a_1\cdot x+\cdots+a_n\cdot x^n,$$

with coefficients in $\mathfrak{A}$ let $H_p(v_0)$ be the following formula of $\mathrm{L}(\mathfrak{A})$,

$$H_p(v_0)=T_n(\bar{a}_0,\bar{a}_1,\ldots,\bar{a}_n,v_0,0),$$

where T_n is defined as in example 5.3.20 and $\bar{a}_0,\bar{a}_1,\ldots,\bar{a}_n$ are the constants of $\mathrm{L}(\mathfrak{A})$ corresponding to the elements $a_0,\ldots,a_n$ of A. $H_p(v_0)$ says '$p(v_0)=0$'. Thus the set of sentences

$$\Delta=\{\neg H_p(c): p \text{ is a polynomial with coefficients in } A\}$$

says that c is a transcendental element over $\mathfrak{A}$. Thus any model, say $\mathfrak{C}$, of $\Sigma_{\mathrm{CF}}\cup D\cup\Delta$ is an algebraically closed field which is (up to isomorphism) an extension of $\mathfrak{A}$ of degree of transcendence at least one over $\mathfrak{A}$. It follows that there is a subfield of $\mathfrak{C}$ which is isomorphic to $\mathfrak{B}$ under a correspondence which leaves the elements of $\mathfrak{A}$ fixed. Thus there is a solution of $(*)$ in $\mathfrak{C}$ or in other words

$$\mathfrak{C}\models(\exists v_0)\ldots(\exists v_m)\,\psi(\bar{a}_{i_0},\ldots,\bar{a}_{i_s},v_0,\ldots,v_m).$$

Let σ be the sentence

$$(\exists v_0)\ldots(\exists v_m)\,\psi(\bar{a}_{i_0},\ldots,\bar{a}_{i_s},v_0,\ldots,v_m)$$

of $\mathrm{L}(\mathfrak{A})$. We have thus shown that any model of $\Sigma_{\mathrm{CF}}\cup D\cup\Delta$ is a model of σ. Therefore by the completeness theorem

$$\Sigma_{\mathrm{CF}}\cup D\cup\Delta\vdash\sigma.$$

Hence, by the finiteness theorem, there is a finite subset

$$\Delta_0=\{\neg H_{p_1}(c),\ldots,\neg H_{p_j}(c)\}$$

of Δ, such that

$$\Sigma_{\mathrm{CF}}\cup D\cup\Delta_0\vdash\sigma. \tag{1}$$

Let $p_0(x)=p_1(x)\cdot p_2(x)\cdot\cdots\cdot p_n(x)$. Then the conjunction of the sentences in Δ_0 is equivalent to $\neg H_{p_0}(c)$. Therefore, by (1) and the deduction theorem,

$$\Sigma_{\mathrm{CF}}\cup D\vdash\neg H_{p_0}(c)\to\sigma.$$

Since c does not occur in $\Sigma_{\mathrm{CF}}\cup D$, it follows that

$$\Sigma_{\mathrm{CF}}\cup D\vdash(\forall v_0)[\neg H_{p_0}(v_0)\to\sigma],$$

and therefore that

$$\Sigma_{\mathrm{CF}}\cup D\vdash(\exists v_0)\neg H_{p_0}(v_0)\to\sigma.$$

The sentence $(\exists v_0)\neg H_{p_0}(v_0)$ says that there is an element which is not a root of $p_0(x)=0$. Since this equation has only finitely many roots and any algebraically closed field is infinite $(\exists v_0)\neg H_{p_0}(v_0)$ holds in $\mathfrak{A}$. But $(\mathfrak{A},a^*)$ is a model of $\Sigma_{\mathrm{CF}}\cup D$, where a^* is an enumeration of A, and therefore

$$(\mathfrak{A},a^*)\models\sigma$$

whence, clearly,

$$\mathfrak{A}\models_a\phi,$$

which is what we desired to prove. This completes the proof in the special case.

In the general case, let $\mathfrak{A}^*$ be the algebraic closure of the extension field $\mathfrak{A}[b_1,\ldots,b_m]$ of $\mathfrak{A}$. The degree of transcendence of $\mathfrak{A}^*$ over $\mathfrak{A}$ is at most m and so there is a chain of $l+1$ algebraically closed fields

$$\mathfrak{A}=\mathfrak{A}_0\subseteq\mathfrak{A}_1\subseteq\cdots\subseteq\mathfrak{A}_l=\mathfrak{A}^*\quad\text{with } l\leqslant m,$$

such that $\mathfrak{A}_{i+1}$ is of degree of transcendence one over $\mathfrak{A}_i$ for $0\leqslant i\leqslant l-1$.

Since $b_1,\ldots,b_m$ are elements of $\mathfrak{A}^*$, $\mathfrak{A}^*\models_a\phi$, hence by what we have just proved $\mathfrak{A}_{l-1}\models_a\phi$, whence $\mathfrak{A}_{l-2}\models_a\phi,\ldots$, whence $\mathfrak{A}\models_a\phi$. This completes the

proof that Σ_{CF} is model complete. However Σ_{CF} is not complete since there are algebraically closed fields of different characteristics and so none of the sentences C_p or their negations is deducible from Σ_{CF}.

For each prime p, let

$$\Sigma_{CF}^{p}=\Sigma_{CF}\cup\{C_p\}$$

and let

$$\Sigma_{CF}^{0}=\Sigma_{CF}\cup\{\neg C_p: p \text{ is a prime}\}.$$

Thus Σ_{CF}^{p} is the theory of algebraically closed fields of characteristic p ($p=0$ or a prime). Σ_{CF}^{p} is model complete since it is an extension of Σ_{CF}.

Let $\mathfrak{A}_p$ be the algebraic closure of the prime field of characteristic p. $\mathfrak{A}_p$ is a prime model of Σ_{CF}^{p}. This gives another proof that Σ_{CF}^{p} is complete.

§ 6. Historical and bibliographical remarks

Example 1.3 is a famous result due to CANTOR [1895]. The remark about the decidability of recursively axiomatizable complete theories is due to JANICZAK [1950]. Vaught's test is due to ŁOŚ [1954] and VAUGHT [1954] where corollaries 1.10, 1.11, 1.13, 1.15 and 1.16 and the places where they were originally proved can be found. Lindenbaum's theorem is due to Adolf Lindenbaum (see TARSKI [1930i]).

The notions of model completeness and of the diagram of a model are both due to Abraham Robinson (see e.g. ROBINSON [1963]). Theorem 4.1 and its consequences are due to KOCHEN [1961]. Robinson's model completeness test was first proved in ROBINSON [1956]. The application of this test to densely and discretely ordered sets is taken from ROBINSON and ZAKON [1960] and that to algebraically closed fields from ROBINSON [1956].

CHAPTER 10

HOMOGENEOUS AND UNIVERSAL STRUCTURES

This chapter is devoted to describing an important algebraic construction. In the next chapter we relate this to the ultraproduct construction and the ideas we use will find applications in ch. 12.

§1. The main definitions

Let $\langle R, < \rangle$ be the ordered set consisting of the rationals together with their usual ordering. Notice that $\langle R, < \rangle$ has the following properties:

(1). If $\{a_1,\ldots,a_k\}$ and $\{b_1,\ldots,b_k\}$ are two sets of elements of R with

$$a_1 < a_2 < \cdots < a_k$$

and

$$b_1 < b_2 < \cdots < b_k,$$

there is an order automorphism h of $\langle R, < \rangle$ which for $1 \leqslant i \leqslant k$, maps a_i onto b_i.

(2). If $\langle X, < \rangle$ is any countable, totally ordered set, there is an embedding of $\langle X, < \rangle$ into $\langle R, < \rangle$.

(1) is obvious and (2) is proved by using the method we adopted to prove the result of example 9.1.3. In addition the rationals are the unique ordered set with these properties in the sense that

(3). If $\langle S, < \rangle$ is any countable totally ordered set with properties (1) and (2) then $\langle S, < \rangle$ is order isomorphic to $\langle R, < \rangle$.

We prove (3) as follows. By (2) S must be infinite. Suppose S has a first element a and let b be any element of S with $a<b$. There is no order automorphism of $\langle S, < \rangle$ mapping a onto b since such an automorphism must leave a fixed. This contradicts (1), so we can conclude that S has no first element. Similarly S has no last element.

Next we show that S is densely ordered. Suppose not and that a,b are elements of S with $a<b$ which have no element of S between them. S has

no last element and so there is some c with $a<b<c$. Then

$$a<b$$

and

$$a<c$$

but there is no order automorphism of S mapping a onto itself and b onto c, since b could not be the image of any element under such an automorphism. This contradicts (1).

Therefore $\langle S,<\rangle$ is a countable densely ordered set without first or last element and so by example 9.1.4 is isomorphic to $\langle R,<\rangle$.

(1) says that the rationals are in a certain sense homogeneous since two finite sets with the same cardinal cannot be distinguished. (2) says that among the countable totally ordered sets they form a universal system and (3) says that any such homogeneous universal system is unique among countable totally ordered sets.

In this chapter we shall be interested in discovering whether other collections of relational structures contain structures with properties analogous to (1), (2) and (3). We shall need a number of definitions which we now list.

Let $\mathscr{M}$ be some collection of relational structures, which we assume to be all of the same type, and let κ be some infinite cardinal. We use $\mathscr{M}_\kappa$ to denote the collection of relational structures in $\mathscr{M}$ of cardinal κ, and $\mathscr{M}_{<\kappa}$ to denote the collection of those structures in $\mathscr{M}$ of cardinal $<\kappa$.

If $\mathfrak{A}\in\mathscr{M}$, we say that $\mathfrak{A}$ is $(\mathscr{M},\kappa)$-*homogeneous* if given any two elements of $\mathscr{M}_{<\kappa}$, say $\mathfrak{B}$ and $\mathfrak{B}'$, which are substructures of $\mathfrak{A}$, and an isomorphism f of $\mathfrak{B}$ onto $\mathfrak{B}'$, there is an automorphism g of $\mathfrak{A}$ which is an extension of f.

We can represent this situation by the commutative diagram 1.a.

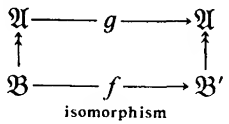

Diagram 1.a

Note that in all diagrams in this chapter double headed arrows indicate injection maps.

$\mathfrak{A}$ is said to be $\mathscr{M}$-*homogeneous* if $\mathfrak{A}$ is of cardinal κ and is $(\mathscr{M},\kappa)$-homogeneous.

The structure $\mathfrak{A}\in\mathscr{M}$ is said to be $(\mathscr{M},\kappa)$-*universal*, if given any structure $\mathfrak{B}\in\mathscr{M}_{<\kappa}$, there is an embedding of $\mathfrak{B}$ into $\mathfrak{A}$. In particular, we say that $\mathfrak{A}$ is $\mathscr{M}$-*universal*, if $\mathfrak{A}$ is $(\mathscr{M},\kappa^+)$-universal, where κ is the cardinal of $\mathfrak{A}$.

Thus $\mathfrak{A}$ is $\mathscr{M}$-universal if any relational structure in $\mathscr{M}$ whose cardinal does not exceed that of $\mathfrak{A}$ can be embedded in $\mathfrak{A}$.

It can be seen that our remarks above show that if $\mathscr{M}$ is the class of totally ordered sets, then $\langle R, < \rangle$ is $\mathscr{M}$-homogeneous and $\mathscr{M}$-universal. Hausdorff showed over fifty years ago that, assuming the G.C.H., for each regular cardinal $\aleph_\alpha$, $\mathscr{M}$ contains a structure of cardinal $\aleph_\alpha$ which is $\mathscr{M}$-homogeneous and $\mathscr{M}$-universal. Such a set is called an η_α-*set*. For more about such sets see HAUSDORFF [1914], GILLMAN and JERISON [1960] and the next chapter.

It also turns out that η_α-sets are unique up to isomorphism and that the condition of $\mathscr{M}$-homogeneity is essential for this to be the case. For example, if η is the order type of the rationals then ordered sets of order types η and $\eta+1$ are both $\mathscr{M}$-universal, but they are not isomorphic.

We now describe a set of conditions on a class $\mathscr{M}$ of relational structures sufficient to guarantee the existence of $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures in $\mathscr{M}$. We say that $\mathscr{M}$ is a κ-*class*, where κ is an infinite cardinal, if $\mathscr{M}$ satisfies the following conditions:

I. $\mathscr{M}$ contains structures of arbitrarily large cardinality.

II. If $\mathfrak{A} \in \mathscr{M}$, and $\mathfrak{A} \simeq \mathfrak{B}$, then $\mathfrak{B} \in \mathscr{M}$.

III. If $\mathfrak{A}, \mathfrak{B} \in \mathscr{M}$, there is some structure $\mathfrak{C} \in \mathscr{M}$ and embeddings f of $\mathfrak{A}$ into $\mathfrak{C}$, and g of $\mathfrak{B}$ into $\mathfrak{C}$.

IV. *The amalgamation property*. If $\mathfrak{A}, \mathfrak{B}_0, \mathfrak{B}_1 \in \mathscr{M}$, f_0 is an embedding of $\mathfrak{A}$ into $\mathfrak{B}_0$ and f_1 is an embedding of $\mathfrak{A}$ into $\mathfrak{B}_1$, then there is some structure $\mathfrak{C} \in \mathscr{M}$ and embeddings g_0 of $\mathfrak{B}_0$ into $\mathfrak{C}$ and g_1 of $\mathfrak{B}_1$ into $\mathfrak{C}$ such that diagram 1.b is commutative:

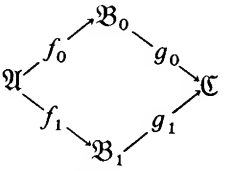

Diagram 1.b

V. *Union of chains*. The union of any chain of relational structures each of which is in $\mathscr{M}$ is itself in $\mathscr{M}$.

VI$_\kappa$. If $\mathfrak{A} \in \mathscr{M}$ and $\mathfrak{C}$ is a substructure of $\mathfrak{A}$ of cardinal $< \kappa$ then there is some $\mathfrak{B} \in \mathscr{M}_{<\kappa}$ such that

$$\mathfrak{C} \subseteq \mathfrak{B} \subseteq \mathfrak{A}.$$

VII$_\kappa$. Each structure in $\mathscr{M}$ has less than κ relations.

We shall be particularly interested in the case when $\mathscr{M}$ is a Δ-elementary class. In this case $\mathscr{M}$ obviously satisfies I and II, and it is usually not difficult to check whether or not $\mathscr{M}$ satisfies III,V,VI$_\kappa$ and VII$_\kappa$. Indeed theorem 4.3.1 shows that if $\mathscr{M} \in EC_\Delta$ then VII$_\kappa$ implies VI$_{\kappa^+}$. On the other hand, no satisfactory criterion is known for deciding whether an arbitrary Δ-elementary class satisfies the amalgamation property. It is known, for example, that the classes of all groups, commutative groups, fields, partially ordered sets, lattices and Boolean algebras have the amalgamation property. Constructions for the first two cases may be found in NEUMANN [1954]. The classes of rings and semi-groups on the other hand, do not have this property.

§ 2. The uniqueness of $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures

This section contains some preliminary results for our discussion of the existence of $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures.

LEMMA 2.1. *If $\mathscr{M}$ is a κ-class, then $\mathscr{M}$ has the following property:*

IV′. *If $\mathfrak{B} \in \mathscr{M}$, $\mathfrak{A}, \mathfrak{A}'$ are substructures of $\mathfrak{B}$, which are also in $\mathscr{M}$, and f is an isomorphism of $\mathfrak{A}$ onto $\mathfrak{A}'$, then there is some $\mathfrak{C} \in \mathscr{M}$ with $\mathfrak{B} \subseteq \mathfrak{C}$ and an embedding g of $\mathfrak{B}$ into $\mathfrak{C}$ which is an extension of f.*

PROOF. Let i, i' be the injection maps of $\mathfrak{A}, \mathfrak{A}'$ into $\mathfrak{B}$, respectively. Then i and $i' \cdot f$ are embeddings of $\mathfrak{A}$ into $\mathfrak{B}$ and so by IV there is a structure $\mathfrak{C} \in \mathscr{M}$ and embeddings g, h of $\mathfrak{B}$ into $\mathfrak{C}$ such that diagram 2.a is commutative:

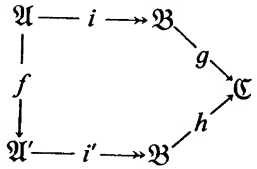

Diagram 2.a

Since, by II, $\mathscr{M}$ is closed under isomorphism, we can choose $\mathfrak{C}$ and h so that $\mathfrak{B} \subseteq \mathfrak{C}$ and h is the injection map of $\mathfrak{B}$ into $\mathfrak{C}$. Since diagram 2.a is commutative, for any $a \in A$, $f(a) = g(a)$ and so g is an extension of f.

Recall that the cardinals $\beth_\xi$ were defined in ch. 0, by $\beth_0 = \aleph_0$, $\beth_{\xi+1} = 2^{\beth_\xi}$, and if λ is a limit ordinal

$$\beth_\lambda = \sum_{\xi < \lambda} \beth_\xi,$$

and that we say that a cardinal κ is a *beth number* if $\kappa = \beth_\xi$ for some ordinal ξ and that κ is a *limit beth number* if $\kappa = \beth_\delta$, where δ is a limit ordinal.

LEMMA 2.2. *Suppose $\mathscr{M}$ is a κ-class and $2^{<\kappa} = \kappa$, then the set of isomorphism types of structures in $\mathscr{M}_{<\kappa}$ is of cardinal at most κ. In particular this is true if either κ is a limit beth number or the G.C.H. holds.*

PROOF. Suppose $\mu \in \omega^\gamma$ is the type of the structures in $\mathscr{M}$, with by VII_κ, $\gamma < \kappa$. Then for each cardinal $\lambda < \kappa$ there are at most $(2^\lambda)^\gamma$ isomorphism types of structures of cardinal $< \kappa$, since in a set of cardinal λ a finitary relation can be chosen in 2^λ ways. Therefore there are at most

$$\sum_{\lambda<\kappa} 2^{\lambda\gamma} = \sum_{\lambda<\kappa} 2^\lambda = 2^{<\kappa}$$

isomorphism types of structures in $\mathscr{M}_{<\kappa}$. Now if $\kappa = \beth_\delta$ with δ a limit ordinal,

$$\kappa = \sum_{\xi<\delta} \beth_\xi = \sum_{\xi<\delta} \beth_{\xi+1} = \sum_{\xi<\delta} 2^{\beth_\xi} = 2^{<\kappa},$$

and the G.C.H. also implies that $\kappa = 2^{<\kappa}$. This completes the proof.

This is the only occasion we shall need to use condition VII_κ directly. We could widen our definition of a κ-class and insist that instead of VII_κ such a class should always satisfy the property described in the lemma above, but it would, in general, be more difficult to check that this new condition held.

A chain $\langle \mathfrak{A}_\xi : \xi < \alpha \rangle$ will be said to be *continuous* if for each limit ordinal $\lambda < \alpha$,

$$\mathfrak{A}_\lambda = \bigcup_{\xi<\lambda} \mathfrak{A}_\xi.$$

LEMMA 2.3. *Let $\mathscr{M}$ be a κ-class and suppose $\mathfrak{A} \in \mathscr{M}_\kappa$. Then $\mathfrak{A}$ is the union of a continuous chain of structures from $\mathscr{M}_{<\kappa}$.*

PROOF. Let λ be the cofinality of κ. Then we can find a chain of substructures of $\mathfrak{A}$ each of cardinal $< \kappa$, say $\langle \mathfrak{A}_\xi : \xi < \lambda \rangle$ such that

$$\mathfrak{A} = \bigcup_{\xi<\lambda} \mathfrak{A}_\xi.$$

We now define the sequence $\langle \mathfrak{B}_\xi : \xi < \lambda \rangle$ by recursion as follows. Each $\mathfrak{B}_\xi$ will be a substructure of $\mathfrak{A}$ which is in $\mathscr{M}_{<\kappa}$.

Suppose $\xi < \lambda$ and for $\zeta < \xi$ we have defined $\mathfrak{B}_\zeta$. Then

$$X = \bigcup_{\zeta<\xi} B_\zeta \cup A_\xi$$

is a subset of A of cardinal $< \kappa$. Therefore, by VI_κ there is a structure $\mathfrak{B}_\xi$

in $\mathscr{M}_{<\kappa}$ such that

$$\mathfrak{A} \restriction X \subseteq \mathfrak{B}_\xi \subseteq \mathfrak{A}.$$

To obtain a continuous chain we define the sequence $\langle \mathfrak{B}'_\xi : \xi < \lambda \rangle$ thus

$$\mathfrak{B}'_0 = \mathfrak{B}_0$$
$$\mathfrak{B}'_\xi = \mathfrak{B}_\eta, \text{ if } \xi = \eta + 1 \text{ is a successor ordinal, } \xi < \lambda,$$
$$\mathfrak{B}'_\delta = \bigcup_{\xi < \delta} \mathfrak{B}_\xi, \text{ if } \delta < \lambda \text{ is a limit ordinal.}$$

By V, $\mathfrak{B}'_\delta$ is in $\mathscr{M}_\kappa$. It is now clear that $\langle \mathfrak{B}'_\xi : \xi < \lambda \rangle$ is a continuous chain of structures from $\mathscr{M}_{<\kappa}$ whose union is $\mathfrak{A}$.

LEMMA 2.4. *If $\mathscr{M}$ is a κ-class, then for each cardinal $\lambda > \kappa$, $\mathscr{M}$ is also a λ-class.*

PROOF. Suppose $\mathscr{M}$ is a κ-class. It will be sufficient to prove that for each $\lambda > \kappa$, $\mathscr{M}$ satisfies VI_λ. We do this by transfinite induction.

Suppose $\lambda > \kappa$, and for all cardinals μ, $\kappa \leqslant \mu < \lambda$, $\mathscr{M}$ satisfies VI_μ. We need to show that $\mathscr{M}$ also satisfies VI_λ.

There are two cases to consider. The first is when λ is a successor cardinal, say $\lambda = \mu^+$. Then, by hypothesis, $\mathscr{M}$ satisfies VI_μ.

Let $\mathfrak{A}$ be in $\mathscr{M}$ and let $\mathfrak{C}$ be a substructure of $\mathfrak{A}$ of cardinal $< \lambda$, and hence of cardinal $\leqslant \mu$. Then

$$\mathfrak{C} = \bigcup_{\xi < \mu} \mathfrak{C}_\xi$$

where $\langle \mathfrak{C}_\xi : \xi < \mu \rangle$ is some chain of substructures of $\mathfrak{C}$ each of cardinal $< \mu$.

We define the chain $\langle \mathfrak{B}_\xi : \xi < \mu \rangle$ of substructures of $\mathfrak{A}$, which are in $\mathscr{M}$, as follows:

$\mathfrak{B}_0$ is any substructure of $\mathfrak{A}$ in $\mathscr{M}$ which is of cardinal $< \mu$. Such a substructure exists since $\mathscr{M}$ satisfies VI_μ.

$\mathfrak{B}_{\xi+1}$ is any substructure of $\mathfrak{A}$ in $\mathscr{M}$ which is an extension of $\mathfrak{A} \restriction (B_\xi \cup C_\xi)$ of cardinal $< \mu$. Again, by VI_μ, such a substructure $\mathfrak{B}_{\xi+1}$ always exists.

If δ is a limit ordinal, $\delta < \mu$, then

$$\mathfrak{B}_\delta = \bigcup_{\xi < \delta} \mathfrak{B}_\xi$$

By V, $\mathfrak{B}_\delta$ is also a substructure of $\mathfrak{A}$ which is in $\mathscr{M}_{<\mu}$. Then

$$\mathfrak{B} = \bigcup_{\xi < \mu} \mathfrak{B}_\xi$$

is a structure in $\mathscr{M}$ of cardinal $\leqslant \mu < \lambda$, satisfying

$$\mathfrak{C} \subseteq \mathfrak{B} \subseteq \mathfrak{A}$$

and so $\mathscr{M}$ satisfies VI_λ.

The second case is when λ is a limit cardinal. In this case it is obvious that $\mathscr{M}$ satisfies VI_λ. This completes the proof of the lemma.

It is convenient at this stage to introduce a new definition, though, mercifully, the next theorem will show that in the case of κ-classes it is not really necessary.

Suppose $\mathfrak{A} \in \mathscr{M}_\kappa$. We say that $\mathfrak{A}$ is *$\mathscr{M}$-homoiogeneous* if whenever $\mathfrak{B},\mathfrak{C} \in \mathscr{M}_{<\kappa}$, with $\mathfrak{B} \subseteq \mathfrak{C} \subseteq \mathfrak{A}$, and f is an embedding of $\mathfrak{B}$ into $\mathfrak{A}$, then there is an embedding g of $\mathfrak{C}$ into $\mathfrak{A}$ which extends f. This definition is due to George Wilmers.

The situation described above is illustrated by the commutative diagram 2.b.

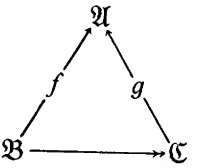

Diagram 2.b

We are now ready for the chief result of this section.

THEOREM 2.5. THE ISOMORPHISM THEOREM

Suppose that $\mathscr{M}$ is a κ-class and $\mathfrak{A},\mathfrak{A}'$ are $\mathscr{M}$ homoiogeneous structures in $\mathscr{M}_\kappa$ such that any substructure of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$ is isomorphic to a substructure of $\mathfrak{A}'$ and vice versa. Then if $\mathfrak{B}$ is a substructure of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$, any embedding of $\mathfrak{B}$ into $\mathfrak{A}'$ can be extended to an isomorphism of $\mathfrak{A}$ onto $\mathfrak{A}'$.

PROOF. First notice that under the conditions of the theorem the following condition holds:

If $\mathfrak{D},\mathfrak{E}'$ are substructures of $\mathfrak{A},\mathfrak{A}'$ respectively which are in $\mathscr{M}_{<\kappa}$ and g is an embedding of $\mathfrak{D}$ into $\mathfrak{E}'$, there is a substructure $\mathfrak{E}$ of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$ which extends $\mathfrak{D}$ and an isomorphism h of $\mathfrak{E}$ onto $\mathfrak{E}'$ which extends g. [*]

We prove [*] as follows. Let i be the injection of $\mathfrak{D}$ into $\mathfrak{A}$. By hypothesis there is some substructure $\mathfrak{E}^*$ of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$ which is isomorphic to $\mathfrak{E}'$. Let θ be an isomorphism of $\mathfrak{E}'$ onto $\mathfrak{E}^*$. Then $\mathfrak{D}^* = \theta \cdot g\,[\mathfrak{D}]$ is a substructure of $\mathfrak{E}^*$ isomorphic to $\mathfrak{D}$, and $i \cdot g^{-1} \cdot \theta^{-1} \cdot i^*$ is an embedding of $\mathfrak{D}^*$ into $\mathfrak{A}$, where i^* is the injection of $\mathfrak{D}^*$ into $\mathfrak{E}^*$. Therefore since $\mathfrak{A}$ is $\mathscr{M}$-homoiogeneous, there is an embedding f of $\mathfrak{E}^*$ into $\mathfrak{A}$ which extends $i \cdot g^{-1} \cdot \theta^{-1} \cdot i^*$. Let $\mathfrak{E} = f\,[\mathfrak{E}^*]$ and let $h = \theta^{-1} \cdot f^{-1}$; it is clear from diagram 2.c that $\mathfrak{E}$ and h have the desired properties.

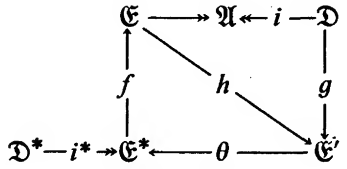

Diagram 2.c

We now set about proving the theorem. Let f be the given embedding of $\mathfrak{B}$ into $\mathfrak{A}$. By lemma 2.3, $\mathfrak{A}$ and $\mathfrak{A}'$ can be written as the union of continuous chains of structures from $\mathscr{M}_{<\kappa}$, say $\mathfrak{A}=\bigcup_{\xi<\lambda}\mathfrak{B}_\xi$ and $\mathfrak{A}'=\bigcup_{\xi<\lambda}\mathfrak{B}'_\xi$, where λ is the cofinality of κ.

We define the chains $\langle\mathfrak{A}_\xi:\xi<\lambda\rangle,\langle\mathfrak{A}'_\xi:\xi<\lambda\rangle$ of substructures of $\mathfrak{A},\mathfrak{A}'$ respectively, from $\mathscr{M}_{<\kappa}$, and for each $\xi<\lambda$ an embedding f_ξ of $\mathfrak{A}_\xi$ into $\mathfrak{A}'_\xi$, by recursion as follows:

Let

$$\mathfrak{A}_0=\mathfrak{B},\quad \mathfrak{A}'_0=f[\mathfrak{B}],$$

and

$$f_0=f.$$

Now suppose $\xi=\zeta+1$ is a successor ordinal, $\xi<\lambda$, and $\mathfrak{A}_\eta,\mathfrak{A}'_\eta,f_\eta$ have been defined for $\eta<\xi$, and in particular for $\eta=\zeta$. By VI_κ there is a substructure $\mathfrak{C}'_\zeta$ of $\mathfrak{A}'$ in $\mathscr{M}_{<\kappa}$ which extends $\mathfrak{A}'\restriction(A'_\zeta\cup B'_\zeta)$. By [*] there is a substructure $\mathfrak{C}_\zeta$ of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$ which extends $\mathfrak{A}_\zeta$ and an isomorphism $f_\zeta{}^*$ of $\mathfrak{C}_\zeta$ onto $\mathfrak{C}'_\zeta$ which extends f_ζ. Again using VI_κ, there is a substructure $\mathfrak{A}_{\zeta+1}$ of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$ which extends $\mathfrak{A}\restriction(C_\zeta\cup B_\zeta)$, and by [*] there is a substructure $\mathfrak{A}'_{\zeta+1}$ of $\mathfrak{A}'$ in $\mathscr{M}_{<\kappa}$ which extends $\mathfrak{C}'_\zeta$ and an isomorphism $f_{\zeta+1}$ of $\mathfrak{A}_{\zeta+1}$ onto $\mathfrak{A}'_{\zeta+1}$ which extends $f_\zeta{}^*$. Diagram 2.d will make the situation clearer:

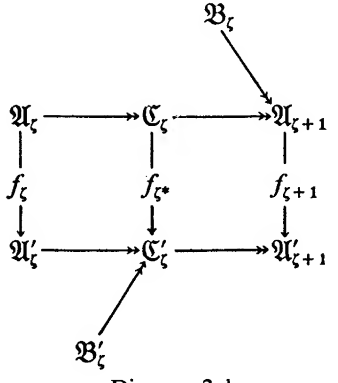

Diagram 2.d

Thus we have defined $\mathfrak{A}_{\zeta+1},\mathfrak{A}'_{\zeta+1}$ and $f_{\zeta+1}$.

Finally we consider the case when $\delta<\lambda$ is a limit ordinal and $\mathfrak{A}_\zeta, \mathfrak{B}_\zeta$ and f_ζ have been defined for all $\zeta<\delta$. In this case we let

$$\mathfrak{A}_\delta=\bigcup_{\zeta<\delta}\mathfrak{A}_\zeta \quad \text{and} \quad \mathfrak{A}'_\delta=\bigcup_{\zeta<\delta}\mathfrak{A}'_\zeta.$$

By V, $\mathfrak{A}_\delta, \mathfrak{A}'_\delta$ are in $\mathcal{M}$ and they are clearly of cardinal $<\kappa$, and we let

$$f_\delta=\bigcup_{\zeta<\delta} f_\zeta.$$

It is clear from our construction that

$$\mathfrak{A}=\bigcup_{\xi<\lambda}\mathfrak{A}_\xi, \quad \mathfrak{A}'=\bigcup_{\xi<\lambda}\mathfrak{A}'_\xi$$

and $h=\bigcup_{\xi<\lambda} f_\xi$ is an isomorphism between $\mathfrak{A}$ and $\mathfrak{A}'$ which extends f. This completes the proof.

The isomorphism theorem has two important consequences.

COROLLARY 2.6. *Suppose $\mathcal{M}$ is a κ-class and $\mathfrak{A}\in\mathcal{M}_\kappa$ then $\mathfrak{A}$ is $\mathcal{M}$-homogeneous iff $\mathfrak{A}$ is $\mathcal{M}$-homoiogeneous.*

PROOF. If $\mathfrak{A}$ is $\mathcal{M}$-homogeneous it is clear that $\mathfrak{A}$ is also $\mathcal{M}$-homoiogeneous.

Now suppose that $\mathfrak{A}$ is $\mathcal{M}$-homoiogeneous; to see that $\mathfrak{A}$ is also $\mathcal{M}$-homogeneous it is sufficient to put $\mathfrak{A}'=\mathfrak{A}$ in theorem 2.5.

COROLLARY 2.7. *If $\mathcal{M}$ is a κ-class, any two $\mathcal{M}$-homogeneous $\mathcal{M}$-universal elements of $\mathcal{M}_\kappa$ are isomorphic.*

PROOF. Suppose $\mathfrak{A},\mathfrak{A}'$ are $\mathcal{M}$-homogeneous, $\mathcal{M}$-universal structures in $\mathcal{M}_\kappa$. By the previous corollary $\mathfrak{A},\mathfrak{A}'$ are both $\mathcal{M}$-homoiogeneous. Since $\mathfrak{A}'$ is $\mathcal{M}$-universal, any substructure of $\mathfrak{A}$ in $\mathcal{M}_{<\kappa}$ is isomorphic to a substructure of $\mathfrak{A}'$ and vice versa. Therefore the conditions of theorem 2.5 apply. By VI_κ, $\mathfrak{A}$ contains a substructure which is in $\mathcal{M}_{<\kappa}$, and by theorem 2.5 the embedding of this structure into $\mathfrak{A}'$, which we know exists, can be extended to an isomorphism of $\mathfrak{A}$ onto $\mathfrak{A}'$. This completes the proof.

Notice that this corollary is the analogue of property (3) of §1 which we observed was possessed by the rationals.

§3. The existence of $\mathcal{M}$-homogeneous, $\mathcal{M}$-universal structures

We are now ready to set about showing that any κ-class $\mathcal{M}$ contains $\mathcal{M}$-homogeneous, $\mathcal{M}$-universal structures. The first step is to show that in order

to find an $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structure of cardinal κ it is sufficient to find a structure which is $\mathscr{M}$-homogeneous, $(\mathscr{M},\kappa)$-universal and of cardinal κ.

THEOREM 3.1. *Let $\mathscr{M}$ be a κ-class. If $\mathfrak{A}$ is an $\mathscr{M}$-homogeneous, $(\mathscr{M},\kappa)$-universal structure of cardinal κ then $\mathfrak{A}$ is also $\mathscr{M}$-universal.*

PROOF. We need to show that $\mathfrak{A}$ is $(\mathscr{M},\kappa^+)$-universal. Suppose $\mathfrak{A}'$ is an element of $\mathscr{M}_{<\kappa^+}$. By lemma 2.3,

$$\mathfrak{A}' = \bigcup_{\xi<\lambda} \mathfrak{A}'_\xi,$$

where λ is the cofinality of κ, and $\langle \mathfrak{A}'_\xi : \xi<\lambda\rangle$ is a continuous chain of structures from $\mathscr{M}_{<\kappa}$.

Using a method similar to that used in the proof of the isomorphism theorem we construct a chain $\langle \mathfrak{A}_\xi : \xi<\lambda\rangle$ of substructures of $\mathfrak{A}$ in $\mathscr{M}_{<\kappa}$ and embeddings f_ξ of $\mathfrak{A}'_\xi$ into $\mathfrak{A}_\xi$.

$\mathfrak{A}$ is $(\mathscr{M},\kappa)$-universal and so there is an embedding f_0 of $\mathfrak{A}'_0$ into $\mathfrak{A}$. We let $\mathfrak{A}_0 = f_0[\mathfrak{A}'_0]$.

Suppose $\xi=\zeta+1$ is a successor ordinal, $\xi<\lambda$, and $\mathfrak{A}_\zeta, f_\zeta$ have been defined. Then, as in the proof of the isomorphism theorem, we can find an embedding $f_{\zeta+1}$ of $\mathfrak{A}'_{\zeta+1}$ into $\mathfrak{A}$ which extends f_ζ. We let $\mathfrak{A}_\xi = f_\xi[\mathfrak{A}'_\xi]$.

Secondly, if $\delta<\lambda$ is a limit ordinal, then since the chain $\langle \mathfrak{A}'_\xi : \xi<\lambda\rangle$ is continuous $\mathfrak{A}'_\delta = \bigcup_{\xi<\delta} \mathfrak{A}'_\xi$, it is sufficient to put $f_\delta = \bigcup_{\xi<\delta} f_\xi$ and $\mathfrak{A}_\delta = \bigcup_{\xi<\delta} \mathfrak{A}_\xi$.

Then $f = \bigcup_{\xi<\lambda} f_\xi$ is an embedding of $\mathfrak{A}'$ into $\mathfrak{A}$. This shows that $\mathfrak{A}$ is $(\mathscr{M},\kappa^+)$-universal, that is, $\mathfrak{A}$ is $\mathscr{M}$-universal.

THEOREM 3.2. *Let $\mathscr{M}$ be a κ-class and suppose $\kappa = 2^{<\kappa}$. Then $\mathscr{M}$ contains an $(\mathscr{M},\kappa)$-universal structure of cardinal κ.*

PROOF. By lemma 2.2, $\mathscr{M}_{<\kappa}$ contains at most κ structures of different isomorphism types. Let $\langle \mathfrak{A}_\xi : \xi<\kappa\rangle$ be a sequence of elements of $\mathscr{M}_{<\kappa}$ containing at least one structure of each of these isomorphism types. We define the chain $\langle \mathfrak{B}_\xi : \xi<\kappa\rangle$ of elements of $\mathscr{M}_{<\kappa^+}$ as follows:

By lemma 2.4, $\mathscr{M}$ satisfies VI_{κ^+} and so, using I, $\mathscr{M}$ contains a structure of cardinal κ, say $\mathfrak{B}_0$.

Suppose $\xi=\zeta+1$ is a successor ordinal, $\xi<\kappa$, and $\mathfrak{B}_\zeta$ has already been defined. By III, $\mathscr{M}$ contains a structure $\mathfrak{C}_\xi$ such that there are embeddings f of $\mathfrak{A}_\zeta$ into $\mathfrak{C}_\xi$ and g of $\mathfrak{B}_\zeta$ into $\mathfrak{C}_\xi$. By II we can clearly take $\mathfrak{C}_\xi$ to be an extension of $\mathfrak{B}_\zeta$ and g to be the injection map. $\mathscr{M}$ satisfies VI_{κ^+} and so there

is some substructure $\mathfrak{B}_\xi$ of $\mathfrak{C}_\xi$ of cardinal κ in $\mathscr{M}$ such that

$$\mathfrak{C}_\xi \restriction (f[A_\zeta] \cup B_\zeta) \subseteq \mathfrak{B}_\xi \subseteq \mathfrak{C}_\xi,$$

We can represent this situation by diagram 3.a.

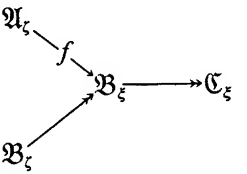

Diagram 3.a

Finally, if $\delta < \kappa$ is a limit ordinal we let

$$\mathfrak{B}_\delta = \bigcup_{\xi<\delta} \mathfrak{B}_\xi.$$

By V, $\mathfrak{B}_\delta \in \mathscr{M}_\kappa$.

Then it is easily seen that $\mathfrak{B} = \bigcup_{\xi<\kappa} \mathfrak{B}_\xi$ is an $(\mathscr{M},\kappa)$-universal structure of cardinal κ.

We shall need the next result from set theory.

LEMMA 3.3. *If $\kappa = \kappa^{<\kappa}$, there are at most κ functions from a subset of κ of cardinal $<\kappa$ into κ. In particular this is true if κ is regular and either is a limit beth number or the G.C.H. holds.*

PROOF. For $\lambda < \kappa$ there are κ^λ functions from a set of cardinal λ into κ and at most κ^λ subsets of κ of cardinal λ. Hence there are at most $\kappa^\lambda \cdot \kappa^\lambda = \kappa^\lambda$ functions from a subset of κ of cardinal λ into κ and thus altogether at most

$$\sum_{\lambda<\kappa} \kappa^\lambda = \kappa^{<\kappa}$$

functions from subsets of κ of cardinal $<\kappa$ into κ.

Now if κ is regular and either κ is a limit beth number or the G.C.H. holds, $\kappa = \kappa^{<\kappa}$ (see BACHMANN [1967] p. 157) and this completes the proof.

Now we can complete our construction by means of

THEOREM 3.4. THE HOMOGENIZATION THEOREM

Let $\mathscr{M}$ be a κ-class and suppose that κ is a regular beth number. Then any structure in $\mathscr{M}$ of cardinal κ has an $\mathscr{M}$-homogeneous extension of cardinal κ which is also in $\mathscr{M}$.

PROOF. Let $\mathfrak{A}_0$ be an element of $\mathscr{M}_\kappa$. By the previous lemma there are

not more than κ isomorphisms between substructures of $\mathfrak{A}_0$ which are in $\mathscr{M}_{<\kappa}$. Let $\langle f_\xi : \xi < \kappa \rangle$ be an enumeration of these isomorphisms, possibly with repetitions.

We construct an extension $\mathfrak{A}_1$ of $\mathfrak{A}_0$ of cardinal κ, which is in $\mathscr{M}$, such that any isomorphism between substructures of $\mathfrak{A}_0$ in $\mathscr{M}_{<\kappa}$ can be extended to an embedding of $\mathfrak{A}_0$ into $\mathfrak{A}_1$. As usual we do this by defining a chain $\langle \mathfrak{A}'_\xi : \xi < \kappa \rangle$ of relational structures:

$$\mathfrak{A}'_0 = \mathfrak{A}_0 .$$

Suppose $\xi = \zeta + 1$ is a successor ordinal, $\xi < \kappa$, and $\mathfrak{A}'_\zeta$ has already been defined. Since $\mathscr{M}$ is a κ-class, condition IV′ of lemma 2.1 holds and therefore there is an extension $\mathfrak{A}'_\xi$ of $\mathfrak{A}'_\zeta$ in $\mathscr{M}$ and an embedding g_ξ of $\mathfrak{A}'_\zeta$ (and *a fortiori* of $\mathfrak{A}_0$) into $\mathfrak{A}'_\xi$ which extends f_ξ. We also know, by lemma 2.4, that $\mathscr{M}$ satisfies VI_{κ^+}, and so we can assume that $\mathfrak{A}'_\xi$ is of cardinal κ.

If $\delta < \kappa$ is a limit ordinal,

$$\mathfrak{A}'_\delta = \bigcup_{\xi<\delta} \mathfrak{A}'_\xi .$$

It is easily seen that $\mathfrak{A}_1 = \bigcup_{\xi<\kappa} \mathfrak{A}'_\xi$ is an extension of $\mathfrak{A}_0$ in $\mathscr{M}_\kappa$ such that any isomorphism between substructures of $\mathfrak{A}_0$ which are in $\mathscr{M}_{<\kappa}$ can be extended to an embedding of $\mathfrak{A}_0$ into $\mathfrak{A}_1$.

We now define the chain $\langle \mathfrak{A}_\xi : \xi < \kappa \rangle$ of structures from $\mathscr{M}_\kappa$ as follows: $\mathfrak{A}_0$ and $\mathfrak{A}_1$ have already been defined.

If $\xi = \zeta + 1 < \kappa$, then $\mathfrak{A}_\xi$ is the structure obtained from $\mathfrak{A}_\zeta$ in exactly the same way that we obtained $\mathfrak{A}_1$ from $\mathfrak{A}_0$.

If $\delta < \kappa$ is a limit ordinal

$$\mathfrak{A}_\delta = \bigcup_{\xi<\delta} \mathfrak{A}_\xi .$$

Now let

$$\mathfrak{A} = \bigcup_{\xi<\kappa} \mathfrak{A}_\xi .$$

$\mathfrak{A}$ is an element of $\mathscr{M}_\kappa$.

Suppose $\mathfrak{B}, \mathfrak{C}$ are substructures of $\mathfrak{A}$ which are in $\mathscr{M}_{<\kappa}$, $\mathfrak{B} \subseteq \mathfrak{C}$ and f is an embedding of $\mathfrak{B}$ into $\mathfrak{A}$. Since κ is regular and card $(C \cup f[B]) < \kappa$, there is some $\xi < \kappa$ such that $\mathfrak{A}_\xi$ is an extension of both $\mathfrak{C}$ and $f[\mathfrak{B}]$.

f is an isomorphism between substructures of $\mathfrak{A}_\xi$ which are in $\mathscr{M}_{<\kappa}$, and so by construction can be extended to an embedding of $\mathfrak{A}_\xi$ into $\mathfrak{A}_{\xi+1}$. Thus f can be extended to an embedding of $\mathfrak{C}$ into $\mathfrak{A}$.

This shows that $\mathfrak{A}$ is $\mathscr{M}$-homoiogeneous. It follows at once from corollary 2.6 that $\mathfrak{A}$ is $\mathscr{M}$-homogeneous.

We can now prove the main theorem of this chapter.

THEOREM 3.5. JÓNSSON'S THEOREM
If $\mathscr{M}$ is a κ-class and $\kappa = \kappa^{<\kappa}$, $\mathscr{M}$ contains an $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structure of cardinal κ which is unique up to isomorphism. In particular this is true if κ is regular and either κ is a limit beth number or the G.C.H. holds.

PROOF. Clearly if $\kappa = \kappa^{<\kappa}$, $\kappa = 2^{<\kappa}$, and so by theorem 3.2, $\mathscr{M}$ contains an $(\mathscr{M},\kappa)$-universal structure of cardinal κ. The homogenization theorem implies that this structure has an $\mathscr{M}$-homogeneous extension of cardinal κ, say $\mathfrak{A}$; clearly this extension is also $(\mathscr{M},\kappa)$-universal. Theorem 3.1 tells us that $\mathfrak{A}$ is also $\mathscr{M}$-universal. Corollary 2.7 to the isomorphism theorem implies that $\mathfrak{A}$ is unique up to isomorphism.

COROLLARY 3.6. (G.C.H.) *If κ is a regular cardinal and $\mathscr{M}$ is a κ-class, $\mathscr{M}$ contains an $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structure of cardinal κ, which is unique up to isomorphism.*

Corollary 3.6 cannot be extended to singular cardinals, as is shown by the following example due to JÓNSSON [1956]. Let $\mathscr{M}$ be the class of all totally ordered sets. It is easy to see that $\mathscr{M}$ is a κ-class for all κ. But if λ is singular, there is no $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structure of cardinal λ. For suppose $\mathfrak{A}$ is such a structure with domain $\{a_\xi : \xi < \lambda\}$. Let $b_0 = a_0$. Now suppose that for $\xi < \zeta$, with $\zeta < \lambda$, b_ξ has already been defined. Since $\mathfrak{A}$ is $\mathscr{M}$-universal it has a subset, say $\{c_\xi : \xi < \zeta + 1\}$, of order type $\zeta + 1$. Since $\mathfrak{A}$ is $\mathscr{M}$-homogeneous there is an automorphism π of $\mathfrak{A}$ with $\pi(c_\xi) = b_\xi$, for $\xi < \zeta$. Let b_ζ be the larger of $\pi(c_\zeta)$ and a_ζ. Then $\{b_\xi : \xi < \lambda\}$ is a cofinal subset of $\mathfrak{A}$ of order type λ. It follows that $\mathfrak{A}$ has a cofinal subset, say X, of order type $\mathrm{cf}(\lambda)$. $\mathrm{cf}(\lambda) + 1 < \lambda$, and so, since $\mathfrak{A}$ is $\mathscr{M}$-universal, it also contains a subset, say Y, of order type $\mathrm{cf}(\lambda) + 1$. Let f be the embedding which sends X onto the initial segment of Y of order type $\mathrm{cf}(\lambda)$. This embedding cannot be extended to an automorphism of $\mathfrak{A}$ (which element could be mapped onto the last element of Y?). This contradicts the assumption that $\mathfrak{A}$ is $\mathscr{M}$-homogeneous.

However it is possible to obtain an extension of corollary 3.6 to singular cardinals by introducing the notion of an *$\mathscr{M}$-special* structure. We do not discuss this here. The original definition and results may be found in MORLEY and VAUGHT [1962]. In the next chapter we explain a similar notion due to CHANG and KEISLER [1966].

§ 4. Full expansions

The notions of this chapter have so far all been of an algebraic nature. We now describe a model theoretic construction which helps us to relate them to model theoretic concepts.

Let $\mathfrak{A}$ be a relational structure of type μ, and let L be the appropriate language for structures of this type. Corresponding to each formula $\phi(v_0,\ldots,v_{n-1})$ of L, with free variables $v_0,\ldots,v_{n-1}$, there is an n-ary relation R_ϕ on A, given by

$$\langle a_0,\ldots,a_{n-1}\rangle \in R_\phi \quad \text{iff} \quad \mathfrak{A} \models \phi[a_0,\ldots,a_{n-1}].$$

The *full expansion* of $\mathfrak{A}$, which we denote by $\mathfrak{A}^*$, is the structure obtained from $\mathfrak{A}$ by adding all the relations R_ϕ, thus

$$\mathfrak{A}^* = (\mathfrak{A}, \{R_\phi : \phi \text{ is a formula of L}\}).$$

All full expansions of structures of type μ are of the same type, say μ^*. Let L* be the appropriate language for structures of this type. For each formula $\phi(v_0,\ldots,v_{n-1})$ with free variables $v_0,\ldots,v_{n-1}$, there is an n-ary predicate letter P of L*, such that given any structure $\mathfrak{A}$ of type μ,

$$\mathfrak{A}^* \models (\forall v_0)\ldots(\forall v_{n-1})[\phi(v_0,\ldots,v_{n-1}) \leftrightarrow P(v_0,\ldots,v_{n-1})].$$

The next lemma is an immediate consequence of this easily verified remark.

LEMMA 4.1. *If $\mathfrak{A},\mathfrak{B}$ are any two relational structures of type μ, then*

(1). $\mathfrak{A} \simeq \mathfrak{B}$ *iff* $\mathfrak{A}^* \simeq \mathfrak{B}^*$.

(2). $\mathfrak{A} \prec \mathfrak{B}$ *iff* $\mathfrak{A}^* \prec \mathfrak{B}^*$ *iff* $\mathfrak{A}^* \subseteq \mathfrak{B}^*$.

(3). $f : \mathfrak{A} \to \mathfrak{B}$ *is an elementary embedding*
iff $f : \mathfrak{A}^* \to \mathfrak{B}^*$ *is an elementary embedding*
iff $f : \mathfrak{A}^* \to \mathfrak{B}^*$ *is an embedding.*

PROOF. This is left as an exercise for the reader.

It can be seen from this that, when we are dealing with full expansions, the notions of extension and elementary extension, and of embedding and elementary embedding coincide. For many purposes it is more convenient to consider full expansions rather than general relational structures. For our purposes the interest of the notion is given by the next theorem. If $\mathscr{M}$ is a class of relational structures, we let $\mathscr{M}^*$ be the collection of full expansions of the structures in $\mathscr{M}$ together with all substructures of these full expansions.

THEOREM 4.2. *Let Σ be a complete set of sentences of* L *and let $\mathscr{M}$ be the collection of all models of Σ. If $\mathscr{M}$ contains an infinite structure and κ is an infinite cardinal $\geqslant$ card(Σ), then $\mathscr{M}^*$ is a κ^+-class.*

PROOF. We check the conditions one by one:

I. If $\mathscr{M}$ contains an infinite structure then by the upward Löwenheim–Skolem theorem $\mathscr{M}$ has models of arbitrarily large cardinality and hence so too does $\mathscr{M}^*$.

II. Trivially $\mathscr{M}$ and hence also $\mathscr{M}^*$ is closed under isomorphism.

III. Suppose $\mathfrak{A}', \mathfrak{B}' \in \mathscr{M}^*$. Then for some models $\mathfrak{A}, \mathfrak{B}$ of Σ, $\mathfrak{A}' \subseteq \mathfrak{A}^*$ and $\mathfrak{B}' \subseteq \mathfrak{B}^*$. Now $\mathfrak{A} \equiv \mathfrak{B}$ since Σ is complete. Therefore, by Frayne's lemma, $\mathfrak{B}$ can be elementarily embedded in some ultrapower $\mathfrak{A}^I/F$ of $\mathfrak{A}$. $\mathfrak{A}^I/F$ is in $\mathscr{M}$ since it is elementarily equivalent to $\mathfrak{A}$. It is now sufficient to observe that $\mathfrak{A}^{*I}/F \simeq (\mathfrak{A}^I/F)^*$; and so $(\mathfrak{A}^I/F)^*$ is an element of $\mathscr{M}^*$ into which both $\mathfrak{A}^*$ and $\mathfrak{B}^*$, and hence both $\mathfrak{A}'$ and $\mathfrak{B}'$, can be embedded.

IV. Suppose $\mathfrak{A}', \mathfrak{B}'_0, \mathfrak{B}'_1 \in \mathscr{M}^*$ and f_0, f_1 are embeddings of $\mathfrak{A}'$ into $\mathfrak{B}'_0, \mathfrak{B}'_1$, respectively. For some models $\mathfrak{B}_0, \mathfrak{B}_1$ of Σ, $\mathfrak{B}'_0 \subseteq \mathfrak{B}_0^*$ and $\mathfrak{B}'_1 \subseteq \mathfrak{B}_1^*$. f_0 and f_1 are also embeddings of $\mathfrak{A}'$ into $\mathfrak{B}_0^*$ and $\mathfrak{B}_1^*$, respectively. Hence if $\{a_\xi : \xi < \alpha\}$ is an enumeration of A', $\mathfrak{B}_0^* \restriction \{f_0(a_\xi) : \xi < \alpha\}$ is isomorphic to $\mathfrak{B}_1^* \restriction \{f_1(a_\xi) : \xi < \alpha\}$ under the natural correspondence. $\mathfrak{B}_0^*, \mathfrak{B}_1^*$ are full expansions and so

$$\mathfrak{B}_0^{*\prime} = (\mathfrak{B}_0^*, \langle f_0(a_\xi) : \xi < \alpha \rangle) \equiv (\mathfrak{B}_1^*, \langle f_1(a_\xi) : \xi < \alpha \rangle) = \mathfrak{B}_1^{*\prime}.$$

Hence again using Frayne's lemma, there is a structure

$$\mathfrak{C}^{*\prime} = (\mathfrak{C}^*, \langle b_\xi : \xi < \alpha \rangle)$$

and elementary embeddings $g_0 : \mathfrak{B}_0^{*\prime} \to \mathfrak{C}^{*\prime}$ and $g_1 : \mathfrak{B}_1^{*\prime} \to \mathfrak{C}^{*\prime}$. It follows from this that diagram 4.a is commutative, and clearly $\mathfrak{C}^* \in \mathscr{M}^*$.

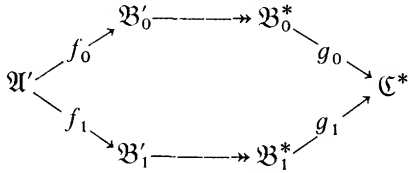

Diagram 4.a

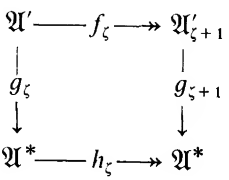

Diagram 4.b

V. Let $\langle \mathfrak{A}'_\xi : \xi < \alpha \rangle$ be a chain in $\mathscr{M}^*$. Suppose $\beta < \alpha$ and we have constructed the chain $\langle \mathfrak{A}_\xi^* : \xi < \beta \rangle$ of full expansions in $\mathscr{M}^*$ so that for $\xi < \beta$ there is an embedding g_ξ of $\mathfrak{A}'_\xi$ into $\mathfrak{A}_\xi^*$. We consider first the case where β is a successor ordinal, say $\beta = \zeta + 1$. Let f_ζ be the injection map of $\mathfrak{A}'_\zeta$ into $\mathfrak{A}'_{\zeta+1}$. By IV, there is a full expansion $\mathfrak{A}^*_{\zeta+1}$ in $\mathscr{M}^*$ and embeddings $g_{\zeta+1}, h_\zeta$ such that diagram 4.b is commutative. Since $\mathscr{M}^*$ is closed under isomorphism, we can choose h_ζ to be an injection map. It follows that $g_\zeta \subseteq g_{\zeta+1}$.

If β is a limit ordinal, we let $\mathfrak{B}' = \bigcup_{\xi<\beta} \mathfrak{A}'_\xi$, $\mathfrak{B}^* = \bigcup_{\xi<\beta} \mathfrak{A}_\xi^*$ and $g = \bigcup_{\xi<\beta} g_\xi$. $\mathfrak{B}' \subseteq \mathfrak{A}'_\beta$ and so $\mathfrak{B}' \in \mathscr{M}^*$. By theorem 4.2.1, for $\xi < \beta$, $\mathfrak{A}_\xi^* \prec \mathfrak{B}^*$. It follows that $\mathfrak{A}_\xi \equiv \mathfrak{B}$, so $\mathfrak{B}$ is a model of Σ whence $\mathfrak{B}^* \in \mathscr{M}^*$. g is an embedding of

$\mathfrak{B}'$ into $\mathfrak{B}^*$ and so we can repeat the argument above with $\mathfrak{A}'_\xi, \mathfrak{A}^*_\xi, g_\xi$ replaced by $\mathfrak{B}', \mathfrak{B}^*, g$.

Having constructed the chain $\langle \mathfrak{A}^*_\xi : \xi < \alpha \rangle$, we put $\mathfrak{A}^* = \bigcup_{\xi<\alpha} \mathfrak{A}^*_\xi$. $\mathfrak{A}^* \in \mathscr{M}^*$ and $g = \bigcup_{\xi<\alpha} g_\xi$ is an embedding of $\mathfrak{A}' = \bigcup_{\xi<\alpha} \mathfrak{A}'_\xi$ into $\mathfrak{A}^*$. Therefore $\mathfrak{A}'$ is isomorphic to a substructure of $\mathfrak{A}^*$ and so $\mathfrak{A}' \in \mathscr{M}^*$.

VI_{κ^+}. This is trivial since $\mathscr{M}^*$ is closed under substructures.

VII_{κ^+}. This is trivial by the choice of κ.

COROLLARY 4.3. *Let Σ be a complete set of sentences of cardinal ρ. Let $\mathscr{M}$ be the collection of all models of Σ. If $\mathscr{M}$ contains an infinite structure then for all regular limit beth numbers $\kappa \geqslant \rho^+$, $\mathscr{M}^*$ contains an $\mathscr{M}^*$-homogeneous, $\mathscr{M}^*$-universal structure of cardinal κ.*

PROOF. This is an immediate consequence of Jónsson's theorem and the previous result.

COROLLARY 4.4. (G.C.H.) *Assuming the G.C.H. one can omit the hypothesis that κ is a limit beth number in the previous corollary.*

The significance of the results of this section will become apparent in the next chapter where we relate them to the ultraproduct construction.

§ 5. Historical and bibliographical remarks

The notion of an $\mathscr{M}$-universal structure was introduced in JÓNSSON [1956] and that of an $\mathscr{M}$-homogeneous structure in JÓNSSON [1960]. In his 1956 paper Jónsson showed that conditions I–VI_κ are sufficient for the existence of $\mathscr{M}$-universal structures and in his 1960 paper that they also imply the existence of $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures and that such structures are unique up to isomorphism. FRAISSÉ [1954] had already discussed this problem in the special case of countable structures.

Jónsson's results were only for structures with a countable number of relations; they were extended to arbitrary structures by MORLEY and VAUGHT [1962] where condition VII_κ occurs. This last, very important paper is in our opinion written in a style which makes it difficult to read and our account in §§ 1–3 is based on the expository work WILMERS [1967], which contains a much clearer explanation of their results. The notion of a full expansion also comes from MORLEY and VAUGHT [1962] from where the results of §4 are taken.

CHAPTER 11

SATURATED MODELS

We introduce in this chapter the important notion of a saturated model and we show that it is closely related to the ideas of the previous chapter.

§1. Saturated models

We have already described the η_α-sets; we now give another, equally informative, description here.

If $\langle X, < \rangle$ is a totally ordered set and A,B are subsets of X, we write $A<B$ if for all $a\in A$ and $b\in B$, $a<b$. Similarly if $c\in X$, we write $A<c<B$ if $a<c<b$ whenever $a\in A$ and $b\in B$.

Let α be some ordinal. $\langle X, < \rangle$ is said to be an *η_α-set* if whenever A,B are subsets of X each of cardinal $<\aleph_\alpha$, with $A<B$, then there is some $c\in X$, with $A<c<B$.

EXAMPLE 1.1. The rational numbers with their usual ordering form an η_0-set.

EXERCISE 1.2. Do the real numbers with their usual ordering form an η_1-set?

As remarked in the previous chapter HAUSDORFF [1914] showed that η_α-sets have the following properties:

(1). *Uniqueness:* Any two η_α-sets of cardinal $\aleph_\alpha$ are isomorphic.

(2). *Universality:* Any totally ordered set of cardinal $\leqslant\aleph_\alpha$ can be embedded in an η_α-set of cardinal $\aleph_\alpha$.

These properties are reminiscent of those of the $\mathscr{M}$-homogeneous, $\mathscr{M}$-universal structures that we described in ch. 10, and they suggest an alternative way of characterizing such structures. Before we describe this we fix some notation. Throughout our discussion we shall assume that $\mu\in\omega^\rho$ is some fixed type, that all relational structures we mention are of this type, that L is the appropriate language for structures of this type and that for

each ordinal ξ, L_ξ is the language obtained from L by adding the ξ-termed sequence of constants $\langle c_\zeta : \zeta < \xi \rangle$.

Let F be the set of all formulas of L having at most the variable v_0 free, and let F_ξ be the corresponding set of formulas of L_ξ.

Suppose Δ is a set of formulas from F_ξ, $\mathfrak{A}$ is a realization of L and $a \in A^\xi$. We say that Δ is *satisfiable* in $(\mathfrak{A},a)$ if there is some $a_+ \in A$ such that for all $\phi \in \Delta$,

$$(\mathfrak{A},a) \models \phi[a_+].$$

In this case we say a_+ *satisfies* Δ in $(\mathfrak{A},a)$. Δ is *finitely satisfiable* in $(\mathfrak{A},a)$ if each finite subset of Δ is satisfiable in $(\mathfrak{A},a)$.

We are now ready for our main definition. Let α be some cardinal. A relational structure $\mathfrak{A}$ is said to be *α-saturated*, if for each ordinal $\xi < \alpha$, and any $a \in A^\xi$, a set of formulas which is finitely satisfiable in $(\mathfrak{A},a)$ is itself satisfiable in $(\mathfrak{A},a)$. Notice that if $\mathfrak{A}$ is α-saturated, then $\mathfrak{A}$ is β-saturated for all $\beta < \alpha$.

We have at once the following simple facts about α-saturated structures.

LEMMA 1.3. *If $\mathfrak{A}$ is finite then $\mathfrak{A}$ is α-saturated for each cardinal α.*

PROOF. Suppose $A = \{a_1,\ldots,a_n\}$, $\xi < \alpha$, $a \in A^\xi$ and Δ is some set of formulas of F_ξ which is not satisfiable in $(\mathfrak{A},a)$; we will show that Δ is not finitely satisfiable in $(\mathfrak{A},a)$.

Since Δ is not satisfiable in $(\mathfrak{A},a)$, for each $k \leqslant n$, there is some $\phi_k \in \Delta$ such that $(\mathfrak{A},a) \models \neg\phi_k[a_k]$, else a_k would satisfy Δ in $(\mathfrak{A},a)$. But then the finite subset $\{\phi_k : k \leqslant n\}$ of Δ is not satisfiable in $(\mathfrak{A},a)$.

LEMMA 1.4. *If $\mathfrak{A}$ is α-saturated, either $\mathfrak{A}$ is finite or of cardinal $\geqslant \alpha$.*

PROOF. Suppose $\mathfrak{A}$ is a relational structure of cardinal β, with $\aleph_0 \leqslant \beta < \alpha$. Let $a \in A^\beta$ be an enumeration of A. The set $\Delta = \{\neg v_0 = c_\xi : \xi < \beta\}$ is finitely satisfiable in $(\mathfrak{A},a)$, but is not itself satisfiable in $(\mathfrak{A},a)$. Thus $\mathfrak{A}$ is not α-saturated.

In an exactly similar way we can prove

LEMMA 1.5. *Let U be a unary predicate letter of* L *and let $\mathfrak{A}$ be an α-saturated realization of* L. *$U_{\mathfrak{A}}$, the value of U in $\mathfrak{A}$, is either finite or of cardinal $\geqslant \alpha$.*

So far, the only examples of saturated structures have been the finite relational structures which lemma 1.3 shows are trivially α-saturated. The question as yet remains open as to whether there are any other α-saturated structures. The next lemma provides the chief step in the proof that most relational structures have α-saturated elementary extensions.

LEMMA 1.6. *Let α be a cardinal greater than or equal to that of the language* L, *and let $\mathfrak{B}$ be a realization of* L *of cardinal 2^α. There exists an elementary extension $g(\mathfrak{B})$ of $\mathfrak{B}$ of cardinal 2^α such that if $\xi<\alpha^+$, and $a\in B^\xi$, then any set of formulas of F_ξ finitely satisfiable in $(\mathfrak{B},a)$ is satisfiable in $(g(\mathfrak{B}),a)$.*

PROOF. Let $b\in B^{2^\alpha}$ be an enumeration of B and let $\mathscr{S}$ be the collection of all sets Σ of formulas of F_{2^α} of cardinal $\leqslant\alpha$ which are finitely satisfiable in $(\mathfrak{B},b)$.

Clearly card $(\mathscr{S})\leqslant 2^\alpha$. Let $\langle\Sigma_\xi:\xi<2^\alpha\rangle$ be an enumeration of $\mathscr{S}$, and let L′ be the language obtained from L_{2^α} by adding a new set $\{d_\xi:\xi<2^\alpha\}$ of constants.

Let Δ be the set of all sentences of L_{2^α} true in $(\mathfrak{B},b)$, and for each $\xi<2^\alpha$ let

$$\Gamma_\xi=\{\phi(d_\xi):\phi\in\Sigma_\xi\},$$

where $\phi(d_\xi)$ is the sentence of L′ obtained from the formula ϕ by replacing all free occurrences of the variable v_0 by d_ξ. We show that

$$\Theta=\Delta\cup\bigcup_{\xi<2^\alpha}\Gamma_\xi$$

is a consistent set of sentences of L′.

Suppose Θ_0 is a finite set of sentences from Θ. Then there is some finite set $\{\xi_0,\ldots,\xi_k\}$ of ordinals $<2^\alpha$ such that

$$\Theta_0\subseteq\Delta\cup\bigcup_{i=0}^{k}\Gamma_{\xi_i}.$$

By the definition of $\mathscr{S}$, for $0\leqslant i\leqslant k$, the set

$$\Theta_0^i=\{\phi:\phi\in\Sigma_{\xi_i}\text{ and }\phi(d_{\xi_i})\in\Theta_0\}$$

is satisfiable in $(\mathfrak{B},b)$. Let b'_{ξ_i} be an element of B such that for $\phi\in\Theta_0^i$, $(\mathfrak{B},b)\models\phi[b'_{\xi_i}]$, and for $\xi\notin\{\xi_0,\ldots,\xi_k\}$, $\xi<2^\alpha$, let b'_ξ be an arbitrary element of B. Then

$$(\mathfrak{B},b,b'),$$

where $b'=\langle b'_\xi:\xi<2^\alpha\rangle$, is a model of Θ_0.

Since every finite subset of Θ has a model, by the compactness theorem, Θ itself has a model, say

$$(\mathfrak{B}_1,b_1,b'_1).$$

By the downward Löwenheim–Skolem theorem we can assume that card $(B_1)=2^\alpha$ and since $(\mathfrak{B}_1,b_1,b'_1)$ is a model of Δ we can assume that $\mathfrak{B}_1$ is an elementary extension of $\mathfrak{B}$.

Now if we put $g(\mathfrak{B})=\mathfrak{B}_1$, it is easily seen that $g(\mathfrak{B})$ has the desired properties.

THEOREM 1.7. *Let α be a cardinal greater than or equal to that of the language* L. *Then each infinite realization of* L *of cardinal $\leqslant 2^\alpha$ has an α^+-saturated elementary extension of cardinal 2^α.*

PROOF. Let $\mathfrak{A}$ be a realization of L of cardinal $\leqslant 2^\alpha$. We define the sequence $\langle \mathfrak{A}_\xi : \xi < \alpha^+ \rangle$ by recursion as follows:

$\mathfrak{A}_0$ is any elementary extension of $\mathfrak{A}$ of cardinal 2^α.

If $\zeta < \alpha^+$ is a successor ordinal, say $\zeta = \xi + 1$, then we let

$$\mathfrak{A}_\zeta = g(\mathfrak{A}_\xi),$$

where g is the function given by the previous lemma.

If $\lambda < \alpha^+$ is a limit ordinal, $\lambda \neq 0$, then we let

$$\mathfrak{A}_\lambda = \bigcup_{\xi<\lambda} \mathfrak{A}_\xi .$$

By the union of chains theorem, $\mathfrak{A}_\lambda$ is an elementary extension of each $\mathfrak{A}_\xi$, with $\xi < \lambda$.

It follows that $\langle \mathfrak{A}_\xi : \xi < \alpha^+ \rangle$ is an elementary chain and hence, using the union of chains theorem again, if

$$\mathfrak{A}' = \bigcup_{\xi<\alpha^+} \mathfrak{A}_\xi ,$$

then $\mathfrak{A}'$ is an elementary extension of each $\mathfrak{A}_\xi$, with $\xi < \alpha^+$. In particular since $\mathfrak{A}_0 = \mathfrak{A}$, $\mathfrak{A}'$ is an elementary extension of $\mathfrak{A}$. It is clear that $\operatorname{card}(\mathfrak{A}') = 2^\alpha$. We conclude the proof by showing that $\mathfrak{A}'$ is α^+-saturated.

Suppose $\xi < \alpha^+$,

$$a = \langle a_\eta : \eta < \xi \rangle \in A'^{\xi},$$

and Σ is a set of formulas of F_ξ which is finitely satisfiable in $(\mathfrak{A}', a)$. For each $\eta < \xi$, there is some $\xi_\eta < \alpha^+$ such that $a_\eta \in A_{\xi_\eta}$. Let

$$\xi^* = \sup \{\xi_\eta : \eta < \xi\} .$$

Since α^+ is regular and $\xi < \alpha^+$, $\xi^* < \alpha^+$. By the choice of ξ^*,

$$a \in (A_{\xi^*})^\xi .$$

For each finite subset Σ_0 of Σ there is some a'_0 in A' which satisfies Σ_0 in $(\mathfrak{A}', a)$. For some $\xi'_0 < \alpha^+$,

$$a'_0 \in A_{\xi'_0} .$$

Hence if

$$\eta_0 = \max\{\xi_0', \xi^*\},$$

since $\mathfrak{A}_{\eta_0} \prec \mathfrak{A}'$, a_0' satisfies Σ_0 in $(\mathfrak{A}_{\eta_0}, a)$.

Let

$$\eta^* = \sup\{\eta_0 : \Sigma_0 \text{ is a finite subset of } \Sigma\}.$$

Again, since $\operatorname{card}(\Sigma) < \alpha^+$ and α^+ is regular,

$$\eta^* < \alpha^+.$$

By the choice of η^*, Σ is finitely satisfiable in $(\mathfrak{A}_{\eta^*}, a)$. Therefore, by construction, Σ is satisfiable in $(\mathfrak{A}_{\eta^*+1}, a)$. Since $\mathfrak{A}_{\eta^*+1} \prec \mathfrak{A}'$, it follows that Σ is satisfiable in $(\mathfrak{A}', a)$, and our proof is completed.

COROLLARY 1.8. (G.C.H.) *Assuming the generalized continuum hypothesis, for each infinite cardinal α every consistent set of sentences with an infinite model has an α^+-saturated model of cardinal α^+.*

If the language L′ is an extension of the language L and $\mathfrak{A}$ is an α-saturated realization of L′, then it is clear that the reduction of $\mathfrak{A}$ to the language L is also α-saturated. The next result, which we shall need to use in the next chapter, shows that in certain circumstances the same thing is true of restrictions of α-saturated structures.

THEOREM 1.9. *Suppose* L′ *is the language obtained from* L *by adding a unary predicate letter* W, *let* $\mathfrak{A}, \mathfrak{B}$ *be realizations of* L *with* $\mathfrak{B} \prec \mathfrak{A}$, *and let* B *be the domain of* $\mathfrak{B}$. *Then if* $(\mathfrak{A}, B)$ *is an* α*-saturated realization of* L′, $\mathfrak{B}$ *is an* α*-saturated realization of* L.

PROOF. Suppose $\xi < \alpha$, $a \in B^\xi$, and Σ is a set of formulas of F_ξ which is finitely satisfiable in $(\mathfrak{B}, a)$. For each finite subset Δ of Σ, let $\phi_\Delta(v_0)$ be the conjunction of the formulas in Δ. By hypothesis, for each finite subset Δ of Σ, there is some $b_0 \in B$ such that

$$(\mathfrak{B}, a) \models \phi_\Delta[b_0].$$

Therefore, since $\mathfrak{B} \prec \mathfrak{A}$,

$$(\mathfrak{A}, a) \models \phi_\Delta[b_0],$$

whence

$$(\mathfrak{A}, B, a) \models (\exists v_0)[\phi_\Delta(v_0) \wedge W(v_0)].$$

Thus $\Sigma' = \Sigma \cup \{W(v_0)\}$ is finitely satisfiable in $(\mathfrak{A}, B, a)$. But $(\mathfrak{A}, B)$ is α-saturated and $a \in A^\xi$. Hence there is some $b_0 \in A$ which satisfies Σ' in $(\mathfrak{A}, B, a)$.

Since $W(v_0)\in\Sigma'$, $b_0\in B$. Again using the fact that $\mathfrak{B}\prec\mathfrak{A}$, it is easily seen that b_0 satisfies Σ in $\mathfrak{B}$.

This completes the proof that $\mathfrak{B}$ is α-saturated.

§ 2. Ultraproducts and saturated structures

In this section we describe Keisler's method of obtaining saturated structures from ultraproducts. We shall be using the facts about *good ultrafilters* that we established in ch. 6, § 1(f). We assume that L is a language with ρ symbols and that all structures we mention are realizations of L. The main result is

THEOREM 2.1. *Let $\{\mathfrak{A}_i: i\in I\}$ be a collection of relational structures and let F be an α-good, ω-incomplete ultrafilter on I. Then the ultraproduct*

$$\mathfrak{A}=\prod \mathfrak{A}_i/F$$

is α-saturated, provided $\alpha>\rho$.

PROOF. Suppose $\xi<\alpha$, $a\in A^{\xi}$ and Σ is a set of formulas of F_{ξ} which is finitely satisfiable in $(\mathfrak{A},a)$. We need to show that Σ is satisfiable in $(\mathfrak{A},a)$.

Let $\mathfrak{B}=(\mathfrak{A},a)$. For $\zeta<\xi$, $a(\zeta)\in A=\prod A_i/F$. Hence $a(\zeta)=\langle a_i(\zeta): i\in I\rangle/F$, where, for $i\in I$, $a_i(\zeta)$ is some element of A_i. Thus if for $i\in I$, $a_i=\langle a_i(\zeta):\zeta<\xi\rangle$, $\mathfrak{B}$ can be written as an ultraproduct, $\mathfrak{B}=\prod(\mathfrak{A}_i,a_i)/F=\prod \mathfrak{B}_i/F$, say.

Since $\xi,\rho<\alpha$, $\operatorname{card}(\Sigma)<\alpha$, and so $\bar{S}_{\omega}(\Sigma)$ also has cardinal $<\alpha$.

We define the map $g:\bar{S}_{\omega}(\Sigma)\to\mathscr{P}(I)$ by putting

$$g(\Delta)=\{i\in I:\Sigma-\Delta \text{ is satisfiable in } \mathfrak{B}_i\}$$

for each $\Delta\in\bar{S}_{\omega}(\Sigma)$.

Clearly, if $\Delta\subseteq\Delta'$, then $g(\Delta)\subseteq g(\Delta')$ and so g is a monotonic function. For each $\Delta\in\bar{S}_{\omega}(\Sigma)$ let $\phi_{\Delta}(v_0)$ be the conjunction of the formulas in $\Sigma-\Delta$, and let

$$\sigma_{\Delta}=(\exists v_0)\,\phi_{\Delta}(v_0).$$

Since Σ is finitely satisfiable in $\mathfrak{B}$, each sentence σ_{Δ} is true in $\mathfrak{B}$, and therefore, by Łoš's theorem,

$$\{i\in I:\mathfrak{B}_i\vDash\sigma_{\Delta}\}\in F$$

for each $\Delta\in\bar{S}_{\omega}(\Sigma)$. But, obviously,

$$\{i\in I:\mathfrak{B}_i\vDash\sigma_{\Delta}\}\subseteq g(\Delta),$$

and so we can conclude that for $\Delta\in\bar{S}_{\omega}(\Sigma)$,

$$g(\Delta)\in F.$$

F is ω-incomplete, and therefore, by lemma 6.1.13, we can find a descending chain

$$I = X_0 \supseteq X_1 \supseteq X_2 \supseteq \cdots \supseteq X_n \supseteq X_{n+1} \supseteq \cdots$$

of elements of F, whose intersection is empty.

We now define the map $g' : \bar{S}_\omega(\Sigma) \to \mathscr{P}(I)$, by putting

$$g'(\Delta) = g(\Delta) \cap X_n,$$

where n is the number of elements in $\Sigma - \Delta$.

Clearly g' is a monotonic function from $\bar{S}_\omega(\Sigma)$ to F and $g' \leqslant g$. Since F is α-good and $\mathrm{card}(\Sigma) < \alpha$, we can conclude that there is a multiplicative function

$$h : \bar{S}_\omega(\Sigma) \to F,$$

with $h \leqslant g'$.

For $i \in I$, let $n(i)$ be the greatest $n < \omega$, such that $i \in X_n$ and let

$$\Sigma(i) = \{\phi \in \Sigma : i \in h(\Sigma - \{\phi\})\}.$$

$\mathrm{Card}(\Sigma(i)) \leqslant n(i)$, for otherwise there would be a subset Σ_0 of $\Sigma(i)$ of cardinal $n(i)+1$ and then, using the fact that h is a multiplicative function, we would have

$$i \in \bigcap_{\phi \in \Sigma_0} h(\Sigma - \{\phi\}) = h(\Sigma - \Sigma_0) \subseteq g'(\Sigma - \Sigma_0) \subseteq X_{n(i)+1},$$

which contradicts the definition of $n(i)$. In particular it follows that $\Sigma(i)$ is finite, and again using the fact that h is multiplicative, we have

$$i \in \bigcap_{\phi \in \Sigma(i)} h(\Sigma - \{\phi\}) = h(\Sigma - \Sigma(i)).$$

But $h \leqslant g$, and so for all $i \in I$,

$$i \in g(\Sigma - \Sigma(i)),$$

whence, by the definition of g, $\Sigma(i)$ is satisfiable in $\mathfrak{B}_i$.

Choose $b \in \prod B_i$, so that for each $i \in I$, $b(i)$ satisfies $\Sigma(i)$ in $\mathfrak{B}_i$. We claim that b/F satisfies Σ in $\mathfrak{B}$. Indeed for all $\phi \in \Sigma$ we have

$$\{i \in I : \mathfrak{B}_i \models \phi[b(i)]\} \supseteq \{i \in I : \phi \in \Sigma(i)\} = h(\Sigma - \{\phi\}) \in F,$$

whence

$$\{i \in I : \mathfrak{B}_i \models \phi[b(i)]\} \in F,$$

and so, by Łoš's theorem

$$\mathfrak{B} \models \phi[b/F].$$

We have therefore shown that Σ is satisfiable in $\mathfrak{B}$ and the proof of the theorem is finished.

COROLLARY 2.2. (G.C.H.) *Assuming the generalized continuum hypothesis, for each infinite cardinal α any relational structure of cardinal α^+ has an α^+-saturated elementary extension of cardinal α^+.*

PROOF. Let $\mathfrak{A}$ be a relational structure of cardinal α^+. By corollary 6.1.25 there is an α^+-good, ω-incomplete ultrafilter F on α. Therefore, by the previous theorem

$$\mathfrak{B}=\mathfrak{A}^\alpha/F$$

is an α^+-saturated elementary extension of $\mathfrak{A}$. Since

$$\alpha^+ \leqslant \mathrm{card}(B) \leqslant (\alpha^+)^\alpha = \alpha^+,$$

$\mathfrak{B}$ is of cardinal α^+.

We have now seen how two quite different constructions lead to the existence of saturated structures. Although the constructions are different the results of the next section show that they lead to the same saturated model.

§ 3. More about saturated structures

Our first result concerns the uniqueness of α-saturated structures.

THEOREM 3.1. *For all cardinals α, any two elementarily equivalent α-saturated relational structures of cardinal α are isomorphic.*

PROOF. Let $\mathfrak{A},\mathfrak{B}$ be elementarily equivalent α-saturated relational structures of cardinal α. Let

$$a=\langle a_\xi : \xi<\alpha\rangle$$

be an enumeration of A and let

$$b=\langle b_\xi : \xi<\alpha\rangle$$

be an enumeration of B. We define, by recursion, two sequences

$$a'=\langle a'_\xi : \xi<\alpha\rangle,$$
$$b'=\langle b'_\xi : \xi<\alpha\rangle$$

of elements of A,B respectively, such that for all $\xi<\alpha$,

$$(A,a'\restriction\xi)\equiv(\mathfrak{B},b'\restriction\xi) \tag{1}$$

Suppose $\xi<\alpha$ and for all $\zeta<\xi$ we have defined a'_ζ,b'_ζ so that (1) holds.

Let λ be the largest limit ordinal $\leqslant \xi$; there are two cases to consider:

(i). For some $n<\omega$, $\xi=\lambda+2n$. In this case we let

$$a'_\xi = a_{\lambda+n}.$$

Now let Σ be the set of all sentences of $L_{\xi+1}$ which hold in $(\mathfrak{A},a'\restriction \xi+1)$, and let Σ' be the set of those formulas of F_ξ which are obtained from Σ by replacing all occurrences of the constant c_ξ by the variable v_0.

Suppose $\{\phi_0,\ldots,\phi_k\}$ is a finite subset of Σ'. Then

$$(\mathfrak{A},a'\restriction \xi) \models \phi_0 \wedge \cdots \wedge \phi_k[a'_\xi]$$

and so

$$(\mathfrak{A},a'\restriction \xi) \models (\exists v_0)[\phi_0 \wedge \cdots \wedge \phi_k].$$

Therefore since, by hypothesis, (1) holds,

$$(\mathfrak{B},b'\restriction \xi) \models (\exists v_0)[\phi_0 \wedge \cdots \wedge \phi_k],$$

and so there is some element of B which satisfies $\{\phi_0,\ldots,\phi_k\}$ in $(\mathfrak{B},b'\restriction \xi)$. Therefore we have shown that Σ' is finitely satisfiable in $(\mathfrak{B},b'\restriction \xi)$. But $\mathfrak{B}$ is α-saturated and therefore Σ' is satisfiable in $(\mathfrak{B},b'\restriction \xi)$. We let b'_ξ be any element of B which satisfies Σ' in $(\mathfrak{B},b'\restriction \xi)$. It then follows that

$$(\mathfrak{A},a'\restriction \xi+1) \equiv (\mathfrak{B},b'\restriction \xi+1)$$

with respect to the extended language $L_{\xi+1}$.

(ii). For some $n<\omega$, $\xi=\lambda+2n+1$. In this case we let

$$b'_\xi = b_{\lambda+n}$$

and a'_ξ is defined in the same way that we defined b'_ξ in case (i) above.

In this way we define sequences $a'\in A^\alpha$ and $b'\in B^\alpha$ such that (1) holds for all $\xi<\alpha$. Therefore

$$(\mathfrak{A},a') \equiv (\mathfrak{B},b').$$

It is clear from the construction that the range of a' is A and that the range of b' is B. Therefore, by theorem 4.1.12, $\mathfrak{A}$ is isomorphic to $\mathfrak{B}$. This completes the proof.

Notice that we have not assumed in this proof that α is infinite. Thus, together with lemma 1.3, this gives us another proof that finite elementarily equivalent relational structures are isomorphic. Also observe that theorems 2.1 and 3.1 together with corollary 6.1.25 provide an alternative proof of Keisler's ultrapower theorem.

We can now begin to describe the connection between α-saturated

structures and the ideas of the previous chapter. It is convenient first to introduce some new definitions.

Let $\mathfrak{A}$ be a relational structure of cardinal α. We say that $\mathfrak{A}$ is *homogeneous* if given any set $B=\{b_\xi:\xi<\beta\}$ of elements of A with $\beta<\alpha$, any map $f:B\to A$ such that

$$(\mathfrak{A},b)\equiv(\mathfrak{A},f(b))$$

with respect to the extended language L_β, where

$$b=\langle b_\xi:\xi<\beta\rangle$$

and

$$f(b)=\langle f(b_\xi):\xi<\beta\rangle,$$

can be extended to an automorphism of $\mathfrak{A}$. Also we say that $\mathfrak{A}$ is *universal* if any relational structure of cardinal $\leqslant\alpha$ elementarily equivalent to $\mathfrak{A}$ can be elementarily embedded in $\mathfrak{A}$.

It is clear from our remarks about full expansions in the previous chapter that if $\mathscr{M}^*$ is the collection of full expansions of all relational structures elementarily equivalent to $\mathfrak{A}$, then

$\mathfrak{A}$ is homogeneous iff $\mathfrak{A}^*$ is $\mathscr{M}^*$-homogeneous,

and

$\mathfrak{A}$ is universal iff $\mathfrak{A}^*$ is $\mathscr{M}^*$-universal.

This shows that our new definitions are closely related to those of the previous chapter. We now connect them up with the idea of an α-saturated structure.

THEOREM 3.2. *Let $\mathfrak{A}$ be an α-saturated relational structure of cardinal α. Then $\mathfrak{A}$ is homogeneous and universal.*

PROOF. This theorem can be proved by using essentially the same method as that used to prove theorem 3.1. We give a sketch of the proof here.

Suppose $B=\{b_\xi:\xi<\beta\}$ is a subset of A of cardinal $\beta<\alpha$ and f is a map of B into A such that, with the same notation as above

$$(\mathfrak{A},b)\equiv(\mathfrak{A},f(b)).$$

Thus $(\mathfrak{A},b),(\mathfrak{A},f(b))$ are elementarily equivalent, α-saturated, relational structures of cardinal α. Therefore, by the previous theorem, they are isomorphic. Any isomorphism between $(\mathfrak{A},b)$ and $(\mathfrak{A},f(b))$ is an automorphism of $\mathfrak{A}$ which extends f. This shows that $\mathfrak{A}$ is homogeneous.

Secondly, suppose $\mathfrak{A}\equiv\mathfrak{B}$ and $\mathrm{card}(\mathfrak{B})\leqslant\alpha$. Using the same method as

used in proving theorem 3.1 we construct a sequence $a \in A^{\beta}$ such that for $\xi < \beta$

$$(\mathfrak{B}, b \restriction \xi) \equiv (\mathfrak{A}, a \restriction \xi),$$

where $b = \langle b_{\xi} : \xi < \beta \rangle$ is an enumeration of B with $\beta \leqslant \alpha$. It follows that

$$(\mathfrak{B}, b) \equiv (\mathfrak{A}, a)$$

and so the map $f : \mathfrak{B} \to \mathfrak{A}$ given by $f(b_{\xi}) = a_{\xi}$ is an elementary embedding of $\mathfrak{B}$ into $\mathfrak{A}$. This shows that $\mathfrak{A}$ is universal.

The converse of this theorem also holds.

THEOREM 3.3. *Let $\mathfrak{A}$ be a relational structure of cardinal $\alpha \geqslant \rho$ which is homogeneous and universal. Then $\mathfrak{A}$ is α-saturated.*

PROOF. Suppose $\xi < \alpha$, $a \in A^{\xi}$ and Σ is a set of formulas of F_{ξ} which is finitely satisfiable in $(\mathfrak{A}, a)$. We need to show that Σ is satisfiable in $(\mathfrak{A}, a)$.

Let Δ be the set of all sentences of L_{ξ} true in $\mathfrak{A}$. Let L'_{ξ} be the language obtained from L_{ξ} by adding a new constant symbol c, and let Σ' be the set of sentences obtained from the formulas of Σ by replacing all free occurrences of v_0 by c.

Since Σ is finitely satisfiable in $(\mathfrak{A}, a)$ every finite subset of $\Delta \cup \Sigma'$ has a model. Therefore, by the compactness theorem, $\Delta \cup \Sigma'$ itself has a model, and since

$$\operatorname{card}(\Delta \cup \Sigma') \leqslant \alpha,$$

we can assume that this model is of cardinal $\leqslant \alpha$. Say it is

$$(\mathfrak{B}, b, b')$$

where $b \in B^{\xi}$, and for $\eta < \xi$, b_{η} is the value of the constant c_{η}, and $b' \in B$ is the value of c.

Since $(\mathfrak{B}, b)$ is a model of Δ, $\mathfrak{B}$ is elementarily equivalent to $\mathfrak{A}$. Therefore, because $\mathfrak{A}$ is universal, there is an elementary embedding, say f, of $\mathfrak{B}$ into $\mathfrak{A}$.

Since f is an elementary embedding, $(\mathfrak{B}, b) \equiv (\mathfrak{A}, f(b))$, but $(\mathfrak{B}, b)$ is a model of Δ and so also $(\mathfrak{B}, b) \equiv (\mathfrak{A}, a)$. It follows that

$$(\mathfrak{A}, f(b)) \equiv (\mathfrak{A}, a).$$

Now $\mathfrak{A}$ is homogeneous, and therefore we can deduce from this that there is an automorphism of $\mathfrak{A}$, say h, which for $\eta < \xi$ maps $f(b_{\eta})$ onto a_{η}. Let $a' = h \cdot f(b')$. Since h and f are elementary embeddings and $(\mathfrak{B}, b, b')$ is a model of Σ', it is clear that a' satisfies Σ in $(\mathfrak{A}, a)$.

This completes the proof that $\mathfrak{A}$ is α-saturated.

Because of the equivalence exhibited by the last two theorems, we can easily translate the results of the previous chapter into the terminology of this. In this way we obtain the following improvement of corollary 2.2.

THEOREM 3.4. (G.C.H.) *Assuming the generalized continuum hypothesis, for each regular infinite cardinal $\alpha \geqslant \rho^+$ any relational structure of cardinal α has an α-saturated elementary extension of cardinal α.*

PROOF. Immediate from corollary 10.4.4 and the remarks above.

We have already remarked that many of the results of ch. 10 proved under the assumption that $\mathscr{M}$ is a κ-class really depend on less restrictive assumptions. Here we remark on just one consequence of this which we shall need in the next chapter.

THEOREM 3.5. *If $\kappa^{<\kappa} = \kappa \geqslant \rho$, every infinite relational structure of cardinal $\leqslant \kappa$ has an elementary extension of cardinal κ which is homogeneous.*

PROOF. Let $\mathfrak{A}$ be an infinite relational structure of cardinal $\leqslant \kappa$, and let $\mathscr{M}^*$ be the collection of all full expansions of relational structures elementarily equivalent to $\mathfrak{A}$ together with all substructures of these full expansions.

By theorem 4.3.3, $\mathfrak{A}$ has an elementary extension of cardinal κ. Now observe that in proving the homogenization theorem we only used conditions II, IV and VI_{κ^+} and that in proving theorem 10.4.2 we showed that $\mathscr{M}^*$ satisfies these conditions. Hence we can deduce that $\mathfrak{A}^*$ has an $\mathscr{M}^*$-homogeneous extension of cardinal κ and thus that $\mathfrak{A}$ has an elementary extension of cardinal κ which is homogeneous.

COROLLARY 3.6. *Every relational structure of cardinal $\aleph_0$ has an elementary extension of cardinal $\aleph_0$ which is homogeneous.*

PROOF. Since $\aleph_0 = \aleph_0^{<\aleph_0}$ this is immediate from theorem 3.5.

THEOREM 3.7. *Let $\mathfrak{A}$ be a countable relational structure which is the union of an elementary chain of homogeneous relational structures. Then $\mathfrak{A}$ is also homogeneous.*

PROOF. Suppose that $\mathfrak{A}$ is the union of the elementary chain $\langle \mathfrak{A}_\xi : \xi < \alpha \rangle$ where each $\mathfrak{A}_\xi$ is homogeneous.

Suppose that $B = \{b_0, \ldots, b_{m-1}\}$ is a finite subset of A and f is a map from B to A such that

$$(\mathfrak{A}, b_0, \ldots, b_{m-1}) \equiv (\mathfrak{A}, f(b_0), \ldots, f(b_{m-1})). \tag{1}$$

We show that f can be extended to an automorphism of $\mathfrak{A}$.

Let $a=\langle a_n : n<\omega\rangle$ be an enumeration of A. We define, by recursion, two sequences

$$\langle b_{m+k} : k<\omega\rangle,$$
$$\langle f(b_{m+k}) : k<\omega\rangle$$

of elements of A such that for all $k<\omega$:

(1). $b_{m+k}=a_j$, if $k=2j$ is even,

and

$f(b_{m+k})=a_j$, if $k=2j+1$ is odd,

and

(2) $(\mathfrak{A},b_0,\ldots,b_{m-1},b_m,\ldots,b_{m+k})\equiv(\mathfrak{A},f(b_0),\ldots,f(b_{m-1}),f(b_m),\ldots,f(b_{m+k}))$.

Suppose $l<\omega$, and for $k\leqslant l$, b_{m+k} and $f(b_{m+k})$ have been defined so as to satisfy (1) and (2). There are two cases to consider.

First assume $l+1=2j$ is even. In this case there is some $\xi<\alpha$ such that all of $b_0,\ldots, b_{m+l}, f(b_0),\ldots,f(b_{m+l})$ and a_j are in A_ξ. Since $\mathfrak{A}_\xi\prec\mathfrak{A}$ it follows from (2) that

$$(\mathfrak{A}_\xi,b_0,\ldots,b_{m+l})\equiv(\mathfrak{A}_\xi,f(b_0),\ldots,f(b_{m+l})),$$

and therefore, since $\mathfrak{A}_\xi$ is homogeneous, there is an automorphism h of $\mathfrak{A}_\xi$ which extends f. We put

$$b_{m+l+1}=a_j$$

and

$$f(b_{m+l+1})=h(a_j)$$

and (1) and (2) are satisfied with k replaced by $l+1$.

In the second case when $l+1=2j+1$ is odd we repeat this argument except that we put $f(b_{m+l+1})=a_j$ and $b_{m+l+1}=h^{-1}(a_j)$.

Thus the sequences $\langle b_{m+k} : k<\omega\rangle,\langle f(b_{m+k}) : k<\omega\rangle$ are defined for all k. Since they satisfy (1) and (2) it follows that we have succeeded in extending f to an automorphism of $\mathfrak{A}$. This completes the proof.

§4. Special relational structures

In this section we describe how theorem 3.4 can be extended to singular cardinals. We have already raised this question in ch. 10 and the example we gave there shows that theorem 3.4 fails to be true for singular cardinals. However we are able to define a class of structures whose properties are

very similar to those of saturated structures, and which can be proved to exist even at singular cardinalities.

The relational structure $\mathfrak{A}$ of cardinal κ is said to be *special* if for each cardinal $\lambda<\kappa$, there is a λ^+-saturated relational structure $\mathfrak{A}_\lambda$ such that $\langle \mathfrak{A}_\lambda : \lambda \text{ is a cardinal, } \lambda<\kappa\rangle$ is an elementary chain whose union is $\mathfrak{A}$. The chain $\langle \mathfrak{A}_\lambda : \lambda<\kappa\rangle$ is called a *specializing chain* for $\mathfrak{A}$. For notational convenience we shall use λ as a variable ranging over cardinals in this section.

We begin with some trivial consequences of our latest definition.

LEMMA 4.1. *Let $\mathfrak{A}$ be an α-saturated, relational structure of cardinal α, then $\mathfrak{A}$ is special.*

PROOF. For $\lambda<\alpha$, let $\mathfrak{A}_\lambda=\mathfrak{A}$. Then $\langle \mathfrak{A}_\lambda : \lambda<\alpha\rangle$ is a specializing chain for $\mathfrak{A}$.

LEMMA 4.2. *If $\mathfrak{A}$ is of cardinal α^+, for some α, then $\mathfrak{A}$ is α^+-saturated iff $\mathfrak{A}$ is special.*

PROOF. One way round follows at once from lemma 4.1. Now suppose $\mathfrak{A}$ is special. If $\langle \mathfrak{A}_\lambda : \lambda<\alpha^+\rangle$ is a specializing chain for $\mathfrak{A}$, $\mathfrak{A}=\mathfrak{A}_\alpha$ and so is α^+-saturated.

Recall that if α is a cardinal

$$2^{<\alpha}=\sum_{\lambda<\alpha} 2^\lambda.$$

The main theorem on the existence of special structures that we shall use is

THEOREM 4.3. *If $\alpha=2^{<\alpha}$, and α is greater than the cardinal of the language* L, *then any relational structure of cardinal $<\alpha$, has an elementary extension which is a special relational structure of cardinal α.*

PROOF. Suppose first that α is a successor cardinal, say $\alpha=\beta^+$. Then $\alpha\leqslant 2^\beta\leqslant 2^{<\alpha}=\alpha$, and so $2^\beta=\beta^+$. The result is now an immediate consequence of theorem 1.7 and lemma 4.1.

Secondly suppose that α is a limit cardinal. This implies that if $\beta<\alpha$, $\beta^+<\alpha$, and we also have that if $\beta<\alpha$, then $2^\beta\leqslant 2^{<\alpha}=\alpha$.

Now suppose $\mathfrak{A}$ is a realization of L of cardinal $\gamma<\alpha$. We define the elementary chain $\langle \mathfrak{B}_\lambda : \lambda<\alpha\rangle$ as follows:

By theorem 1.7, $\mathfrak{A}$ has a γ^+-saturated elementary extension, say $\mathfrak{A}'$ of cardinal 2^γ. For $\lambda\leqslant\gamma$ we let

$$\mathfrak{B}_\lambda=\mathfrak{A}'.$$

If λ is a successor cardinal, $\gamma<\lambda<\alpha$, say $\lambda=\beta^+$, then we let $\mathfrak{B}_\lambda$ be a β^{++}-saturated elementary extension of $\mathfrak{B}_\beta$ of cardinal 2^{β^+}.

If δ is a limit cardinal, $\gamma<\delta<\alpha$, then we let $\mathfrak{B}_\delta$ be a δ^+-saturated elementary extension of

$$\bigcup_{\lambda<\delta}\mathfrak{B}_\lambda$$

of cardinal 2^δ.

The existence of these $\mathfrak{B}_\lambda$ is guaranteed by theorem 1.7. Then $\langle\mathfrak{B}_\lambda:\lambda<\alpha\rangle$ is a specializing chain for

$$\mathfrak{B}=\bigcup_{\lambda<\alpha}\mathfrak{B}_\lambda.$$

Therefore $\mathfrak{B}$ is a special model, which by the union of chains theorem is an elementary extension of $\mathfrak{A}$.

COROLLARY 4.4. *If α is a limit beth number greater than the cardinal of the language* L, *then any set, Σ, of sentences of* L *with an infinite model has a special model of cardinal α.*

PROOF. If Σ has a model, Σ has a model of cardinal $<\alpha$. Since α is a limit beth number $2^{<\alpha}=\alpha$. The result now follows immediately from the previous theorem.

Analogously to theorem 3.1 we have

THEOREM 4.5. *If $\mathfrak{A}$ and $\mathfrak{B}$ are elementarily equivalent special models of the same cardinal, then $\mathfrak{A}$ is isomorphic to $\mathfrak{B}$.*

PROOF. By corollary 8.1.2, we need only consider the case when $\mathfrak{A}$ and $\mathfrak{B}$ are infinite. Suppose they are both of cardinal α and have specializing chains

$$\langle\mathfrak{A}_\lambda:\lambda<\alpha\rangle$$

and

$$\langle\mathfrak{B}_\lambda:\lambda<\alpha\rangle$$

respectively.

Our proof that $\mathfrak{A}\simeq\mathfrak{B}$ will be an extension of our proof of theorem 3.1.

Let $a\in A^\alpha$ and $b\in B^\alpha$ be enumerations of A and B respectively. Clearly we can choose a so that whenever $\xi<\alpha$ and $\mathrm{card}(\xi)=\lambda$, then $a_\xi\in A_\lambda$, and similarly for b. We are going to define by recursion two sequences $a'\in A^\alpha$, $b'\in B^\alpha$ such that for each $\nu<\alpha$, with $\mathrm{card}(\nu)=\lambda$,

(i). $a'_\nu\in A_\lambda$ and $b'_\nu\in B_\lambda$,

(ii). $(\mathfrak{A},a'\restriction\nu+1)\equiv(\mathfrak{B},b'\restriction\nu+1)$,

and

(iii). $a'_\nu=a_{\eta+n}$, if $\nu=\eta+2n$ with $n<\omega$,
and
$b'_\nu=b_{\eta+n}$, if $\nu=\eta+2n+1$ with $n<\omega$.

Suppose $\mu<\alpha$, card $(\mu)=\lambda$, and (i), (ii) and (iii) hold for all $\nu<\mu$. Then by (ii),

$$(\mathfrak{A},a'\restriction\mu)\equiv(\mathfrak{B},b'\restriction\mu), \tag{1}$$

and by (i)

$$a'\restriction\mu\in(A_\lambda)^\mu$$

and

$$b'\restriction\mu\in(B_\lambda)^\mu. \tag{2}$$

Therefore since $\mathfrak{A}_\lambda\prec\mathfrak{A}$ and $\mathfrak{B}_\lambda\prec\mathfrak{B}$, it follows from (1) and (2) that

$$(\mathfrak{A}_\lambda,a'\restriction\mu)\equiv(\mathfrak{B}_\lambda,b'\restriction\mu). \tag{3}$$

Since $\mathfrak{A}_\lambda$ and $\mathfrak{B}_\lambda$ are both λ^+-saturated and $\mu<\lambda^+$ we can conclude from (3) that we can find a'_μ,b'_μ in A_λ,B_λ, respectively so that (i), (ii) and (iii) are satisfied when ν is replaced by μ, exactly as we did when proving theorem 3.1. This completes the recursive definition of a', b'.

Clearly, by (iii), a',b' are enumerations of A,B respectively, and by (ii)

$$(\mathfrak{A},a')\equiv(\mathfrak{B},b').$$

We can therefore conclude, by theorem 4.1.12, that

$$\mathfrak{A}\simeq\mathfrak{B},$$

thus completing the proof.

It is clear that, as was the case with α-saturated models, reductions of special models are also special. Also it is easy to see that theorem 1.9 still holds if 'α-saturated' is replaced by 'special'.

§ 5. Historical and bibliographical remarks

Saturated relational structures have a complicated history, particularly as they were studied before the concept was isolated and given a name. The reader is referred to KEISLER [1961] for the details. Here we only remark that the idea grew out of some independent work of Morley and Vaught. The theorems in § 1 are mostly due to these logicians. The connection of the notion with the ultraproduct construction is due to KEISLER [1964ii] and the equivalence of saturated and homogeneous universal structures comes from KEISLER [1961] where one half of the implication is attributed to Vaught.

The notion of a special model is due to MORLEY and VAUGHT [1962], although our definition is a simplified version taken from CHANG and KEISLER [1966].

CHAPTER 12

APPLICATIONS OF ULTRAPRODUCTS

We begin this chapter by showing that two classical constructions in model theory, Gödel's construction in his proof of the completeness theorem and Skolem's construction of a countable non-standard model of arithmetic, can both be viewed as ultraproduct constructions. Skolem's result leads on to the more powerful theorem of MacDowell and Specker on proper elementary end extensions of models of arithmetic, and this leads on to a discussion of cardinal-like models and various two cardinal theorems which find an application in the next chapter.

§1. Gödel's completeness proof

In this section we give a version of Gödel's original [1930] proof of the completeness theorem for predicate calculus using the ultraproduct construction to simplify the argument at one point. Our proof will only be for sentences not containing the equality symbol though its extension to arbitrary sentences is quite straightforward.

A sentence is said to be in *Skolem normal form* if it is of the form

$$(\exists x_1)\ldots(\exists x_m)(\forall y_1)\ldots(\forall y_n)\,\phi(x_1,\ldots,x_m,y_1,\ldots,y_n),$$

where $\phi(x_1,\ldots,x_m,y_1,\ldots,y_n)$ is an open formula. We will need to make use of the following well known result.

THEOREM 1.1. SKOLEM'S NORMAL FORM THEOREM
For each sentence σ there is a sentence τ in Skolem normal form, which is provable iff σ is provable, and universally valid iff σ is universally valid.

A proof of this result can be found, for example, in CHURCH [1956], pp. 226 and 230, or MENDELSON [1964], p. 88.

Up to now we have used 'ω' to denote the set of finite ordinals; in this section however we use 'N' to denote the set of natural numbers, which is

really the same thing, and we use 'Ω' to denote the sequence of natural numbers. So $\Omega = \langle 0,1,2,3,\ldots\rangle$.

A simple argument now shows that the completeness theorem holds for universal sentences.

LEMMA 1.2. *If $\sigma = (\forall v_1)\ldots(\forall v_n)\,\phi(v_1,\ldots,v_n)$ is a universal sentence, then either σ is provable or there is a model $\mathfrak{A}$ of $\neg\sigma$, whose domain is the set of natural numbers. Indeed $\mathfrak{A}$ can be chosen so that $\mathfrak{A} \vDash_\Omega \neg\phi(v_1,\ldots,v_n)$, that is, so that $\mathfrak{A} \vDash \neg\phi[1,\ldots,n]$.*

PROOF. If ϕ is a substitution instance of a propositional tautology, $\vdash \phi$, and so, by the rule of Generalization, $\vdash \sigma$.

Thus if σ is not provable, ϕ is not a substitution instance of a tautology. Therefore there is an assignment, say V, of truth values to the atomic formulas occurring in ϕ under which ϕ gets the value *false*.

The realization $\mathfrak{A}$ is defined as follows: the domain of $\mathfrak{A}$ is the set N of natural numbers. If P is a k-ary predicate letter of L its value $P_{\mathfrak{A}}$ in $\mathfrak{A}$ is the k-ary relation given by $\langle n_1,\ldots,n_k\rangle \in P_{\mathfrak{A}}$ iff the atomic formula $P(v_{n_1},\ldots,v_{n_k})$ is assigned the value *true* in the assignment V.

Clearly $\mathfrak{A} \vDash \neg\phi[1,\ldots,n]$, that is, $\mathfrak{A} \vDash_\Omega \neg\phi(v_1,\ldots,v_n)$, and so $\mathfrak{A}$ is a model of $\neg\sigma$. This completes the proof.

THEOREM 1.3. GÖDEL'S COMPLETENESS THEOREM
Every universally valid sentence is a theorem.

PROOF. By Skolem's normal form theorem it is sufficient to consider a sentence, say

$$\sigma = (\exists u_1)\ldots(\exists u_p)(\forall v_1)\ldots(\forall v_q)\,\phi(u_1,\ldots,u_p,v_1,\ldots,v_q),$$

in Skolem normal form.

We shall write this sentence, for abbreviation, as

$$(\exists \boldsymbol{u})(\forall \boldsymbol{v})\,\phi(\boldsymbol{u},\boldsymbol{v}),$$

and adopt similar abbreviations throughout the proof.

Let $\langle \boldsymbol{x}_k : k\in\omega\rangle$ be an enumeration of all p-tuples of variables of L and let $\langle \boldsymbol{y}_k : k\in\omega\rangle$ be a sequence of q-tuples of variables of L such that none of the variables in $\boldsymbol{y}_k$ occur in any $\boldsymbol{y}_i$ with $i<k$, or any $\boldsymbol{x}_i$ with $i\leqslant k$.

For each $k\in\omega$ let

$$\phi_k = \phi(\boldsymbol{x}_k,\boldsymbol{y}_k),$$
$$\psi_k = \phi_1 \vee \cdots \vee \phi_k$$

and let σ_k be the closure of ψ_k.

We prove by induction on k that for all $k\in\omega$, $\sigma_k\vdash\sigma$:

$$\sigma_1=(\forall x_1)(\forall y_1)\,\phi(x_1,y_1)$$

and so, obviously, $\sigma_1\vdash\sigma$. Now suppose $\sigma_{k-1}\vdash\sigma$. None of the variables in y_k occurs in ψ_{k-1} and hence

$$(\forall y_k)[\psi_{k-1}\vee\phi(x_k,y_k)]\vdash\psi_{k-1}\vee(\forall y_k)\,\phi(x_k,y_k),$$

that is

$$(\forall y_k)\,\psi_k\vdash\psi_{k-1}\vee(\forall y_k)\,\phi(x_k,y_k),$$

whence

$$\sigma_k\vdash\psi_{k-1}\vee(\forall y_k)\,\phi(x_k,y_k).$$

Now clearly

$$\vdash(\forall y_k)\,\phi(x_k,y_k)\rightarrow\sigma,$$

and so

$$\sigma_k\vdash\psi_{k-1}\vee\sigma.$$

Hence, by generalization,

$$\sigma_k\vdash\sigma_{k-1}\vee\sigma.$$

But, by hypothesis, $\sigma_{k-1}\vdash\sigma$ and therefore $\sigma_k\vdash\sigma$ which completes our induction.

One of the following two cases holds:

(1). For some $k\in\omega$, $\vdash\sigma_k$. In this case, by what we have just proved, $\vdash\sigma$.

(2). For all $k\in\omega$, not $\vdash\sigma_k$. Then, since σ_k is a universal sentence with matrix ψ_k, it follows from the previous lemma that for each $k\in\omega$ we can define a realization $\mathfrak{A}_k$ of L, whose domain is the set of natural numbers, such that

$$\mathfrak{A}_k\vDash_\Omega\neg\psi_k$$

Suppose $j>k$. Clearly, then, $\vdash\psi_k\rightarrow\psi_j$ and so $\vdash\neg\psi_j\rightarrow\neg\psi_k$. Since, by construction, $\mathfrak{A}_j\vDash_\Omega\neg\psi_j$, it follows that $\mathfrak{A}_j\vDash_\Omega\neg\psi_k$. Thus $\{j\in\omega:\mathfrak{A}_j\vDash\psi_k\}$ is finite. Therefore, if F is a non-principal ultrafilter on ω,

$$\{j\in\omega:\mathfrak{A}_j\vDash_\Omega\psi_k\}\notin F,$$

and hence, by Łoš's theorem, for each $k\in\omega$,

$$\mathfrak{A}\vDash_{d(\Omega)}\neg\psi_k,$$

where $\mathfrak{A}=\prod\mathfrak{A}_k/F$ and d is the canonical embedding of N into N^ω/F.

It follows at once that for each $k\in\omega$,

$$\mathfrak{A}\vDash_{d(\Omega)}\neg\phi_k. \tag{1}$$

Let

$$\mathfrak{B} = \mathfrak{A} \restriction d[N].$$

Since $d(\Omega) \in d[N]^{\omega}$, and $\neg\phi_k$ is an open formula, it follows from (1) that for each $k \in \omega$,

$$\mathfrak{B} \vDash_{d(\Omega)} \neg\phi_k,$$

that is, for each $k \in \omega$,

$$\mathfrak{B} \vDash_{d(\Omega)} \neg\phi(x_k, y_k),$$

whence for each $k \in \omega$,

$$\mathfrak{B} \vDash_{d(\Omega)} \neg(\forall y_k)\, \phi(x_k, y_k).$$

Now as k runs through ω, x_k runs through all p-tuples of variables of L. Also the range of $d(\Omega)$ is the domain of $\mathfrak{B}$. Therefore

$$\mathfrak{B} \vDash (\forall x_k) \neg(\forall y_k)\, \phi(x_k, y_k),$$

and so

$$\mathfrak{B} \vDash \neg\sigma.$$

Thus we have shown that σ is either provable or not universally valid, completing the proof.

It is not difficult to extend this result to sentences containing the equality symbol, but since we have already given another proof of the completeness theorem we leave this as an exercise for the reader.

COROLLARY 1.4. HERBRAND'S THEOREM
σ is provable iff for some $k \in \omega$,

$$\psi_k = \phi(x_1, y_1) \vee \cdots \vee \phi(x_k, y_k)$$

is a substitution instance of a propositional tautology.

PROOF. If ψ_k is a substitution instance of a propositional tautology, $\vdash \psi_k$. Therefore, by generalization, $\vdash \sigma_k$ and so, since $\sigma_k \vdash \sigma$, $\vdash \sigma$.

Conversely, if none of the formulas ψ_k is a substitution instance of a propositional tautology, then, by the proof of lemma 1.2, for each $k \in \omega$ there is a model of $\neg\sigma_k$ and so σ_k is not provable. Therefore, by the proof of theorem 1.3, there is a model of $\neg\sigma$ and so σ is not provable. This completes the proof.

§ 2. A non-standard model of arithmetic

The appropriate language for arithmetic, L, has one binary predicate letter S to represent the successor relation, two ternary predicate letters A and M

to represent addition and multiplication and the constant symbol 0 to represent zero.

The *standard model of arithmetic*, denoted by $\mathcal{N}$, is the model whose domain is the set N of natural numbers, whose relations are the usual ones and in which the value of 0 is 0. By the *theory of complete arithmetic* we mean the set $Th(\mathcal{N})$ of all sentences of L true in $\mathcal{N}$. It follows from Gödel's incompleteness theorem, GÖDEL [1931], that $Th(\mathcal{N})$ cannot be recursively axiomatized. However it is possible to find a recursive set of axioms which enable one to prove a very large collection of true sentences of first order arithmetic. We will not bother to specify such a set of axioms for *formal arithmetic* here. A set of axioms appropriate for the language L can be found in HIGMAN [1963] and other axiomatizations can be found in most of the standard text books. By the *theory of formal arithmetic* we mean the set of sentences provable in some system of formal arithmetic, to be specific say the system described in Higman's lecture notes mentioned above.

It follows from Gödel's incompleteness theorem that there are non-elementarily equivalent, countable models of formal arithmetic and thus models not isomorphic to the standard model. Such models are called *non-standard models of formal arithmetic.*

Also we know that if F is a non-principal ultrafilter on ω then $\mathcal{N}^\omega/F$ is an elementary extension of $\mathcal{N}$ of cardinal $2^{\aleph_0}$ and so is not isomorphic to $\mathcal{N}$. Any model of $Th(\mathcal{N})$ not isomorphic to $\mathcal{N}$ is called a *non-standard model of complete arithmetic* or just a *non-standard model of arithmetic.*

Both these results leave open the possibility that all countable models of complete arithmetic are isomorphic, that is, that $Th(\mathcal{N})$ is $\aleph_0$-categorical. However it was shown by SKOLEM [1934] (see also SKOLEM [1955]) that it is possible to have countable non-standard models of arithmetic. We now describe a version of Skolem's construction, again using ultraproducts.

We have already remarked that if F is a non-principal ultrafilter on ω then $\mathcal{N}^\omega/F$ is an uncountable non-standard model of arithmetic. We obtain a countable non-standard model by taking a suitable elementary substructure of $\mathcal{N}^\omega/F$.

Let d be the canonical embedding of $\mathcal{N}$ into $\mathcal{N}^\omega/F$. Let $R(x,y)$ be the formula

$$(\exists z)[\neg z=0 \wedge A(x,z,y)].$$

It is clear that

$$\mathcal{N} \models R[m,n] \quad \text{iff} \quad m<n$$

in the usual ordering of the natural numbers. For each model of complete

arithmetic $\mathfrak{A}$ we will use $<$ to denote the total ordering of the domain given by

$$a<b \quad \text{iff} \quad \mathfrak{A} \models R[a,b].$$

Since $\mathcal{N}^\omega/F$ is non-standard there is some m^* in N^ω/F such that for all $n\in\omega$, $d(n)<m^*$. Let $X=d[N]\cup\{m^*\}$. By theorem 4.3.1 $\mathcal{N}^\omega/F\restriction X$ has a countable extension which is an elementary substructure of $\mathcal{N}^\omega/F$, say $\mathfrak{A}$. Since $\mathfrak{A}$ contains m^*, it is not isomorphic to $\mathcal{N}$, and so is a countable non-standard model of arithmetic.

In this argument we have not specified the domain of $\mathfrak{A}$ very precisely. In fact it is possible to be more explicit and to construct a countable non-standard model of arithmetic with a known domain. This is what Skolem originally did and we now describe his construction, except that we use ultraproducts instead of Skolem's original combinatorial argument.

A function f of n variables is said to be *definable* in $\mathcal{N}$ if there is a formula $\phi(v_1,\ldots,v_n,v_{n+1})$ of L such that for all $m_1,\ldots,m_n,m\in N$,

$$\mathcal{N} \models \phi[m_1,\ldots,m_n,m] \quad \text{iff} \quad f(m_1,\ldots,m_n)=m.$$

We let $\mathscr{F}$ be the set of all functions of one variable which are definable in $\mathcal{N}$. Since we are identifying N with ω we can regard $\mathscr{F}$ as a subset of N^ω.

LEMMA 2.1. (a) *For each $k\in\omega$, the constant function with value k is in $\mathscr{F}$.*

(b) *The identity function id, where for $k\in\omega$, $id(k)=k$, is in $\mathscr{F}$.*

(c) *If $\Phi(v_0,v_1,\ldots,v_n)$ is a formula with free variables $v_0,\ldots,v_n$ and $f_1,\ldots,f_n$ are all functions in $\mathscr{F}$, then there is some function g in $\mathscr{F}$ such that for all $k\in\omega$, if there is some $l\in\omega$, such that*

$$\mathcal{N} \models \Phi[l,f_1(k),\ldots,f_n(k)],$$

then

$$\mathcal{N} \models \Phi[g(k),f_1(k),\ldots,f_n(k)].$$

PROOF. (a). If $\phi(v_1,v_2)$ is $(\exists!^k v_0)\, R(v_0,v_2)$, where $R(x,y)$ is the formula described above, then

$$\mathcal{N} \models \phi[m,n] \quad \text{iff} \quad n=k.$$

(b). If $\phi(v_1,v_2)$ is $v_1=v_2$ then

$$\mathcal{N} \models \phi[m,n] \quad \text{iff} \quad m=n.$$

(c). For $1\leqslant i\leqslant n$, let $\psi_i(v_1,v_2)$ be the formula which defines f_i in N. We let

g be the function defined as follows:

If for some $l \in \omega$,

$$\mathcal{N} \models \Phi[l, f_1(k), \ldots, f_n(k)],$$

then $g(k)$ is the least such k, otherwise $g(k) = 0$.

Let $\phi(v_1, v_2)$ be the following formula

$$\begin{aligned}(\forall u_1)\ldots(\forall u_n)\{[\psi_1(v_1,u_1) \wedge \cdots \wedge \psi_n(v_1,u_n)] \rightarrow [\{(\exists v_0)\, \Phi(v_0,u_1,\ldots,u_n) \\ \rightarrow [\Phi(v_2,u_1,\ldots,u_n) \wedge (\forall v_3)\{R(v_3,v_2) \rightarrow \neg \Phi(v_3,u_1,\ldots,u_n)\}]\} \\ \wedge \{(\forall v_0) \neg \Phi(v_0,u_1,\ldots,u_n) \rightarrow v_2 = 0\}]\}.\end{aligned}$$

It is easily seen that $\mathcal{N} \models \phi[m,n]$ iff $n = g(m)$. Thus $g \in \mathscr{F}$ and so satisfies the conditions of (c).

Now let F be any non-principal ultrafilter on ω. We let

$$\mathscr{F}/F = \{f/F \in \boldsymbol{N}^\omega/F : f \in \mathscr{F}\},$$

and we let

$$\mathfrak{A} = (\mathcal{N}^\omega/F) \restriction (\mathscr{F}/F).$$

THEOREM 2.2. *$\mathfrak{A}$ is a countable non-standard model of arithmetic which is an elementary extension of $d[\mathcal{N}]$.*

PROOF. Since there are only countably many distinct formulas of L, $\mathscr{F}$ is countable and hence so too is $\mathfrak{A}$. By lemma 2.1(a) $d[\boldsymbol{N}]$ is a subset of the domain of $\mathfrak{A}$. Thus $d[\mathcal{N}] \subseteq \mathfrak{A}$; we will show that $\mathfrak{A} \prec \mathcal{N}^\omega/F$ and hence that $d[\mathcal{N}] \prec \mathfrak{A}$.

Suppose $\Phi(v_0, v_1, \ldots, v_n)$ is a formula of L and $f_1/F, \ldots, f_n/F$ are elements of $d[\boldsymbol{N}]$ such that for some $f/F \in \boldsymbol{N}^\omega/F$,

$$\mathcal{N}^\omega/F \models \Phi[f/F, f_1/F, \ldots, f_n/F].$$

Then, by Łoš's theorem,

$$X = \{k \in \omega : \mathcal{N} \models \Phi[f(k), f_1(k), \ldots, f_n(k)]\} \in F.$$

Let g be the definable function corresponding to the formula Φ and the definable functions $f_1, \ldots, f_n$ given by lemma 2.1(c). Then whenever $k \in X$ there is some $l \in \boldsymbol{N}$, namely $l = f(k)$, such that

$$\mathcal{N} \models \Phi[l, f_1(k), \ldots, f_n(k)].$$

Therefore

$$X \subseteq Y = \{k \in \omega : \mathcal{N} \models \Phi[g(k), f_1(k), \ldots, f_n(k)]\}.$$

Since $X \in F$, $Y \in F$ and so

$$\mathcal{N}^\omega/F \models \Phi[g/F, f_1/F, \ldots, f_n/F].$$

Now $g/F \in \mathscr{F}/F$ and therefore, by corollary 4.1.9,

$$\mathfrak{A} \prec \mathscr{N}^{\omega}/F.$$

Since $d[\mathscr{N}] \prec \mathscr{N}^{\omega}/F$, we can immediately conclude that

$$d[\mathscr{N}] \prec \mathfrak{A}.$$

$\mathfrak{A}$ is therefore, up to isomorphism, a countable elementary extension of the standard model of arithmetic. It is easily seen that id/F, which by lemma 2.1(b) is in $\mathfrak{A}$, is greater than any of the elements in $d[N]$. Thus $\mathfrak{A}$ is not isomorphic to $\mathscr{N}$. This concludes the proof.

Let $\mathfrak{A},\mathfrak{B}$ be any two models of complete arithmetic, with $\mathfrak{B}$ an extension of $\mathfrak{A}$. We say that $\mathfrak{B}$ is a *proper extension* of $\mathfrak{A}$ if $B-A$ is not empty, and that $\mathfrak{B}$ is an *end extension* if for all $a \in A$ and $b \in B-A$, $a<b$, that is, if all the 'new' elements are bigger than all the 'old' elements.

It is clear that all elementary extensions of the standard model of arithmetic are end extensions. Thus theorem 2.2 can be restated as:

the standard model of arithmetic has a *countable proper elementary end extension.*

We show in the next section that all models of formal arithmetic have proper elementary end extensions of the same cardinal.

§ 3. MacDowell and Specker's theorem

For each $n \in \omega$, we can prove, in formal arithmetic, the sentence which says that there is a unique element having exactly n predecessors, namely the sentence

$$(\exists ! x)(\exists !^{n} y)\, R(y,x).$$

It follows that given any model, say $\mathfrak{A}$, of formal arithmetic, for each $n \in \omega$, there is a unique element, say $n_{\mathfrak{A}}$, in the domain of $\mathfrak{A}$ with just n predecessors. We call the elements in $\{n_{\mathfrak{A}} : n \in \omega\}$ the *natural numbers* of $\mathfrak{A}$. It is clear that $\mathfrak{A} \restriction \{n_{\mathfrak{A}} : n \in \omega\}$ is isomorphic to the standard model of arithmetic. Thus each model of formal arithmetic can be regarded as an extension of the standard model. We have already remarked that all elementary extensions of the standard model are end extensions. Clearly $\mathfrak{A}$ is an end extension of $\mathfrak{A} \restriction \{n_{\mathfrak{A}} : n \in \omega\}$. In this section we show that any model of formal arithmetic has a proper elementary end extension of the same cardinality. We need first the following fact about the standard model of arithmetic.

LEMMA 3.1. *Let $f(n_1,\ldots,n_k,n)$ be a function of $k+1$ variables and let X be an infinite set of natural numbers. Then there is an infinite subset Y of X such that for each k-tuple $\langle n_1,\ldots,n_k\rangle$ of natural numbers either*

(a). *there is some number p and some m in Y such that whenever n is in Y and $n>m$, then*

$$f(n_1,\ldots,n_k,n)=p$$

or

(b). *for each number p there is some m in Y such that whenever n is in Y and $n>m$,*

$$f(n_1,\ldots,n_k,n)>p.$$

PROOF. There is some map π mapping the set of natural numbers onto all k-tuples of natural numbers. For $r\in\omega$, if $\pi(r)=\langle r_1,\ldots,r_k\rangle$ we write $f_r(n)$ for $f(r_1,\ldots,r_k,n)$.

We define the descending sequence of infinite sets $\langle X_n:n\in\omega\rangle$ as follows:

$$X_0=X.$$

Suppose X_r has been defined. Since X_r is infinite either there is some infinite subset of X_r on which f_r is constant, in which case we let X_{r+1} be this infinite set, or there is some infinite subset of X_r on which f_r is a strictly increasing function, in this case we let X_{r+1} be this infinite set. Now for $r\in\omega$, we let y_r be the least element of X_r greater than the elements y_s for $s<r$. It is not difficult to see that the set

$$Y=\{y_r:r\in\omega\}$$

has the desired properties.

Let L be the language appropriate for formal arithmetic. If we examine the proof of lemma 3.1 it can be seen that the set Y is defined uniformly from f and X and that the whole argument could be carried out in formal arithmetic. We can conclude from this that we can associate with each formula $\psi(v_0)$ and each formula $\Phi(v_1,\ldots,v_k,x,y)$ which defines a function of $k+1$ variables, say $f(v_1,\ldots,v_k,x)$, another formula $\psi'(v_0)$ such that the following sentences are provable in formal arithmetic:

A(i). $(\forall x)(\exists y)[\psi(y)\wedge x<y]\rightarrow(\forall x)(\exists y)[\psi'(y)\wedge x<y]$.

'If ψ holds for arbitrarily large numbers, so too does ψ'.'

A(ii). $(\forall x)[\psi'(x)\rightarrow\psi(x)]$.

'The numbers for which ψ' holds are a subset of those for which ψ holds.'

A(iii). $(\forall x)(\exists y)[\psi(y) \wedge x < y] \rightarrow$
$(\forall v_1)\ldots(\forall v_k)\{(\exists t)(\exists x)[\psi'(x) \wedge (\forall y)(x < y \wedge \psi'(y) \rightarrow \Phi(v_1,\ldots,v_k,y,t))] \vee$
$(\forall t)(\exists x)[\psi'(x) \wedge (\forall y)(\forall z)(x < y \wedge \psi'(y) \wedge \Phi(v_1,\ldots,v_k,y,z) \rightarrow z > t)]\}$.

'If ψ holds for arbitrarily large numbers, then for all $v_1,\ldots,v_k$ either there is some number t, and some number x for which ψ' holds and such that if $y > x$ and ψ' holds for y, then $f(v_1,\ldots,v_k,y) = t$; or for each number t there is some x for which ψ' holds and such that if $y > x$ and ψ' holds for y then $f(v_1,\ldots,v_k,y) > t$.'

It follows from this that all these sentences are true in any model of formal arithmetic. We use this fact subsequently.

Let $\mathfrak{A}$ be any model of formal arithmetic with domain A and suppose $\text{card}(A) = \mathfrak{m}$. We say that a function $f: A^n \rightarrow A$ is *parametrically definable* in $\mathfrak{A}$ if there is some formula

$$\Phi(u_1,\ldots,u_m,v_1,\ldots,v_n,v_{n+1})$$

of L such that for some $b_1,\ldots,b_m$ in A, for all $a_1,\ldots,a_n,a_{n+1}$ in A,

$$\mathfrak{A} \models \Phi[b_1,\ldots,b_m,a_1,\ldots,a_n,a_{n+1}] \quad \text{iff} \quad f(a_1,\ldots,a_n) = a_{n+1}.$$

In this case we say that f is *defined by Φ with respect to the parameters* $b_1,\ldots,b_m$.

We let $\mathscr{F}$ be the set of all functions from A into A which are parametrically definable in $\mathfrak{A}$. Clearly $\text{card}(\mathscr{F}) = \mathfrak{m}$, and the functions in $\mathscr{F}$ satisfy the analogous properties to those laid down in lemma 2.1. (Note that in proving the analogue of 2.1(c) we use the fact that non-empty *definable* sets have a least element, although in a non-standard model of arithmetic with respect to the natural ordering the domain is not well-ordered.)

It follows from this that if we take an ultrapower $\mathfrak{A}^A/F$ of $\mathfrak{A}$ and then restrict it to $\mathscr{F}/F$, we obtain, up to isomorphism, an elementary extension of $\mathfrak{A}$ of cardinal $\mathfrak{m}$. In order to ensure it is an *end* extension we need to choose F with some care.

L is a countable language and hence we can arrange the formulas of L in a sequence

$$\langle \Phi_k : k \in \omega \rangle.$$

For each $k \in \omega$, and each finite sequence $\boldsymbol{b} = \langle b_1,\ldots,b_m \rangle$ of elements of A, we let $f_{k,\boldsymbol{b}}$ be the function defined by Φ_k with respect to the parameters $b_1,\ldots,b_m$, if there is such a function. Otherwise we let $f_{k,\boldsymbol{b}}$ be the constant function with value 0. Thus for each $f \in \mathscr{F}$ there is some $k \in \omega$, and some finite sequence $\boldsymbol{b}$ of elements of A such that $f_{k,\boldsymbol{b}} = f$.

LEMMA 3.2. *There is a sequence* $\langle X_n : n \in \omega \rangle$ *of subsets of* A *such that*

(1). $A = X_0 \supseteq X_1 \supseteq X_2 \supseteq \cdots \supseteq X_n \supseteq \cdots$,

(2). *each* X_n *contains arbitrarily large elements of* A, *that is for each* $a \in A$ *there is some* $b \in X_n$ *with* $a < b$,

and

(3). *for each* $k \in \omega$ *and each finite sequence* $\boldsymbol{b}$ *of elements of* A *either*

(a). *there is some* $b \in A$ *and some* $a_0 \in X_k$ *such that if* $a \in X_k$ *and* $a_0 < a$, *then* $f_{k,\boldsymbol{b}}(a) = b$,

or

(b). *for each* $b \in A$ *there is some* $a_0 \in X_k$ *such that if* $a \in X_k$ *and* $a_0 < a$, *then* $b < f_{k,\boldsymbol{b}}(a)$.

PROOF. We have already remarked that the formal analogue of lemma 3.1 holds in $\mathfrak{A}$, that is that the sentences A(i)–A(iii) are valid in $\mathfrak{A}$. Lemma 3.2 now follows immediately.

For each $a \in A$ let $X^a = \{b \in A : a < b\}$. It is easily seen that the collection

$$\{X_n : n \in \omega\} \cup \{X^a : a \in A\},$$

where the sets X_n are those given by the previous lemma, has the finite intersection property and so can be extended to an ultrafilter F on A. Let d be the canonical embedding of $\mathfrak{A}$ into $\mathfrak{A}^A/F$. We will show that

$$\mathfrak{B} = \mathfrak{A}^A/F \restriction \mathscr{F}/F$$

is a proper elementary end extension of $d[\mathfrak{A}]$. By the remarks above it will be sufficient to show that $\mathfrak{B}$ is an end extension of $d[\mathfrak{A}]$.

Suppose f/F is in the domain of $\mathfrak{B}$. Then $f \in \mathscr{F}$ and hence for some $k \in \omega$ and some finite sequence $\boldsymbol{b}$ of elements of A, $f = f_{k,\boldsymbol{b}}$.

Either (3a) or (3b) of lemma 3.2 holds. If (3a) holds there is some $b \in A$ and some $a_0 \in X_k$ such that

$$X_k \cap X^{a_0} = \{a \in X_k : a_0 < a\} \subseteq \{a \in A : f_{k,\boldsymbol{b}}(a) = b\}.$$

By construction, $X_k \cap X^{a_0} \in F$, and so

$$\{a \in A : f_{k,\boldsymbol{b}}(a) = b\} \in F.$$

Hence

$$f/F = f_{k,\boldsymbol{b}}/F = d(b),$$

thus $f/F \in d[A]$.

If (3b) holds, then, by a similar argument, for each $b \in A$,

$$f/F > d(b).$$

Therefore either f/F is in $d[A]$ or is greater than all the elements of $d[A]$.

This proves that $\mathfrak{B}$ is an end extension of $d[\mathfrak{A}]$. This completes our proof of

THEOREM 3.3. MACDOWELL AND SPECKER'S THEOREM
Every model of formal arithmetic has a proper elementary end extension of the same cardinality.

We will give an important application of this result in the next section.

§ 4. Cardinal-like orderings

Let $\langle X, < \rangle$ be a totally ordered set. Y is an *initial segment* of X if $Y \subseteq X$ and for all $y \in Y$ and $x \in X$, if $x < y$ then $x \in Y$. $<$ is said to be *a cardinal-like* ordering if each initial segment of X is of smaller cardinal than that of X. In particular if X is of cardinal $\mathfrak{m}$, we say that $<$ is an $\mathfrak{m}$-like ordering. For example an ordered set of order type ω is $\aleph_0$-like, since all its initial segments are finite. It is easily seen that all $\aleph_0$-like orderings are of order type ω. However at higher cardinalities there are more possibilities. Thus if θ is any order type of cardinal $< \aleph_\alpha$, then the order type $\theta \cdot \omega_\alpha$ is $\aleph_\alpha$-like.

Suppose that L is a countable language which contains a binary predicate letter P, and that Σ is a set of sentences of L. We say that Σ *has an* $\mathfrak{m}$*-like model* if Σ has a model of cardinal $\mathfrak{m}$ in which the value of P is an $\mathfrak{m}$-like ordering of the domain. Σ has a *cardinal-like model* if it has an $\mathfrak{m}$-like model for some $\mathfrak{m}$.

A natural question to ask, and one that will be of especial interest in the next chapter, is: if Σ has an $\mathfrak{m}$-like model then for which other cardinals $\mathfrak{n}$ does Σ have an $\mathfrak{n}$-like model? We get some information about this as a consequence of MacDowell and Specker's theorem.

THEOREM 4.1. *If Σ is a set of sentences with an $\aleph_0$-like model, then for each infinite cardinal $\mathfrak{m}$, Σ has an $\mathfrak{m}$-like model.*

PROOF. Let $\mathfrak{A}$ be an $\aleph_0$-like model of Σ. Without loss of generality we can assume that the domain of $\mathfrak{A}$ is the set of natural numbers and that the value of P, say $<$, is their usual ordering. Let L′ be the language obtained from L by adding two new ternary predicate letters A, M, a new binary predicate letter S and a new constant symbol 0.

Let

$$\mathfrak{A}' = (\mathfrak{A}, A', M', S', 0),$$

where A', M', S' are the relations of addition, multiplication and successor on the set of natural numbers. $\mathfrak{A}'$ is therefore the standard model of arithmetic with some extra relations. Since these extra relations do not affect the proof of MacDowell and Specker's theorem, by making use of this

result we can define the sequence $\langle \mathfrak{A}'_\xi : \xi \in \mathfrak{m} \rangle$ of structures as follows:

$$\mathfrak{A}'_0 = \mathfrak{A}',$$

if λ is a limit ordinal $< \mathfrak{m}$,

$$\mathfrak{A}'_\lambda = \bigcup_{\xi<\lambda} \mathfrak{A}'_\xi,$$

and if $\xi = \eta + 1$ is a successor ordinal $< \mathfrak{m}$, $\mathfrak{A}'_\xi$ is the proper elementary end extension of $\mathfrak{A}'_\eta$, of the same cardinal as $\mathfrak{A}'_\eta$, which MacDowell and Specker's theorem tells us exists.

Let

$$\mathfrak{A}'_{\mathfrak{m}} = \bigcup_{\xi<\mathfrak{m}} \mathfrak{A}'_\xi.$$

It is easily seen that $\mathfrak{A}_{\mathfrak{m}}$, the reduction of $\mathfrak{A}'_{\mathfrak{m}}$ to the language L, is an $\mathfrak{m}$-like model of Σ.

It is very important to notice that the above proof only works because $\mathfrak{A}'$ is the *standard model* of arithmetic with some extra relations. The proof of lemma 3.1 is unaffected by the presence of these additional relations, and so the sentences A(i)–A(iii) of the extended language are provable in formal arithmetic. Since the structures $\mathfrak{A}'_\xi$ are all elementarily equivalent to $\mathfrak{A}'$, these sentences hold in all the $\mathfrak{A}'_\xi$ and so MacDowell and Specker's theorem applies to them.

The following example shows that the assumption that MacDowell and Specker's theorem applies to *any* model of formal arithmetic with extra relations leads to a contradiction.

Let $\mathfrak{A}$ be a countable non-standard model of arithmetic and let R be a relation defining a one-one map of the natural numbers of $\mathfrak{A}$ onto the domain of $\mathfrak{A}$. If we assume that MacDowell and Specker's theorem applies to $\mathfrak{A}' = (\mathfrak{A}, R)$, and its elementary extensions, we obtain, as in the proof of theorem 4.1, an uncountable elementary end extension, say $\mathfrak{B}' = (\mathfrak{B}, R')$, of $\mathfrak{A}'$. Since $\mathfrak{A}' \prec \mathfrak{B}'$ and $\mathfrak{B}'$ is an end extension of $\mathfrak{A}'$, R' defines a one–one map between a countable subset of $\mathfrak{B}$ and the domain of $\mathfrak{B}$. But $\mathfrak{B}$ is uncountable and so this is a contradiction.

We owe this remark and counter-example to F. R. Drake.

§ 5. Two-cardinal theorems

Before we give further results about cardinal-like models we need to say something about two-cardinal theorems. Let L be a language containing a unary predicate letter U and let Σ be a set of sentences of L. We say that

Σ *admits* the pair of cardinals $\langle\alpha,\beta\rangle$ if Σ has a model of cardinal α in which the value of U is a set of cardinal β.

We now ask: if Σ admits $\langle\alpha,\beta\rangle$ then for which other pairs of cardinals $\langle\gamma,\delta\rangle$ does Σ admit $\langle\gamma,\delta\rangle$? Our first answer is a simple consequence of the Löwenheim–Skolem theorem. Throughout the rest of this section we shall use $\alpha,\beta,\gamma,\delta$ to range over infinite cardinals.

THEOREM 5.1. *Let Σ be a set of sentences of cardinal δ. If Σ admits $\langle\alpha,\beta\rangle$ then Σ admits $\langle\gamma,\gamma\rangle$ for any $\gamma\geqslant\delta$. Further if $\beta\leqslant\gamma\leqslant\alpha$, then Σ admits $\langle\gamma,\beta\rangle$.*

PROOF. We add to L the set $\langle c_\xi:\xi<\gamma\rangle$ of constants and let

$$\Sigma'=\Sigma\cup\{c_\xi\neq c_\eta:\xi<\eta<\gamma\}\cup\{U(c_\xi):\xi<\gamma\}.$$

Σ admits $\langle\alpha,\beta\rangle$ and so Σ has a model $\mathfrak{A}$ of cardinal α in which $U_{\mathfrak{A}}$, the value of U, is a set of cardinal β. Since α and β are infinite it follows that each finite subset of Σ' has a model and therefore, by the compactness theorem, Σ' itself has a model. The Löwenheim–Skolem theorem implies that Σ', which is of cardinal γ, has a model $\mathfrak{B}$ of cardinal γ. Clearly $\mathrm{card}(U_{\mathfrak{B}})=\gamma$ and so Σ admits $\langle\gamma,\gamma\rangle$.

Now suppose $\beta\leqslant\gamma\leqslant\alpha$. By theorem 4.3.1 there is an elementary substructure $\mathfrak{B}$ of $\mathfrak{A}$ of cardinal γ, which extends $\mathfrak{A}\restriction U_{\mathfrak{A}}$. It follows that $U_{\mathfrak{B}}=U_{\mathfrak{A}}$. $\mathfrak{B}$ is a model of Σ and so Σ admits $\langle\gamma,\beta\rangle$.

The next result is a simple application of the ultraproduct construction.

THEOREM 5.2. *If Σ admits $\langle\alpha,\beta\rangle$ then for any cardinal γ, Σ admits $\langle\alpha^\gamma,\beta^\gamma\rangle$.*

PROOF. Let $\mathfrak{A}$ be a model of Σ of cardinal α such that $\mathrm{card}(U_{\mathfrak{A}})=\beta$. Let F be a regular ultrafilter on γ. Then the ultrapower

$$\mathfrak{B}=\mathfrak{A}^\gamma/F$$

is also a model of Σ, since it is elementarily equivalent to $\mathfrak{A}$. By theorem 6.3.21, $\mathrm{card}(B)=\alpha^\gamma$, and since

$$U_{\mathfrak{B}}=(U_{\mathfrak{A}})^\gamma/F,$$

$\mathrm{card}(U_{\mathfrak{B}})=\beta^\gamma$, by the same result.

This completes the proof.

Recall that if α,β are cardinals,

$$\alpha^{<\beta}=\sum_{\delta<\beta}\alpha^\delta.$$

Our next result is a slightly longer application of ultraproducts.

THEOREM 5.3. *If Σ admits $\langle\alpha,\beta\rangle$, then for any infinite cardinal γ, Σ admits $\langle\alpha^{<\gamma},\beta^{<\gamma}\rangle$.*

PROOF. Let $\mathfrak{A}$ be a model of Σ of cardinal α, in which $U_{\mathfrak{A}}$, the value of U, is of cardinal β.

We define the elementary chain $\langle\mathfrak{A}_\delta:\delta<\gamma\rangle$ by recursion as follows:

$$\mathfrak{A}_\omega=\mathfrak{A}.$$

Suppose μ^+ is an infinite successor cardinal $<\gamma$ and $\mathfrak{A}_\mu$ has already been defined and is an elementary extension of $\mathfrak{A}$ of cardinal $\alpha^{<\mu}$.

Let F_μ be a regular ultrafilter on μ. We let $\mathfrak{A}_{\mu^+}$ be an elementary extension of $\mathfrak{A}_\mu$ which is isomorphic to the ultrapower $\mathfrak{A}_\mu{}^\mu/F_\mu$. Then

$$\begin{aligned}\operatorname{card}(A_{\mu^+})&=\operatorname{card}(A_\mu)^\mu, \quad \text{by theorem 6.3.21}\\ &=(\alpha^{<\mu})^\mu\\ &=\alpha^\mu\\ &=\alpha^{<\mu^+}\end{aligned}$$

Secondly suppose that μ is an infinite limit cardinal $<\gamma$ and for $\delta<\mu$ $\mathfrak{A}_\delta$ has been defined and is an elementary extension of $\mathfrak{A}$ of cardinal $\alpha^{<\delta}$. Then we let

$$\mathfrak{A}_\mu=\bigcup_{\delta<\mu}\mathfrak{A}_\delta.$$

Clearly $\mathfrak{A}_\mu$ is an elementary extension of $\mathfrak{A}$ of cardinal $\alpha^{<\mu}$.

It follows from this that

$$\mathfrak{A}_\gamma=\bigcup_{\delta<\gamma}\mathfrak{A}_\delta$$

is an elementary extension of $\mathfrak{A}$ of cardinal $\alpha^{<\gamma}$. Similarly the value of U in $\mathfrak{A}_\gamma$ can be seen to be a set of cardinal $\beta^{<\gamma}$. Since $\mathfrak{A}_\gamma$, being an elementary extension of $\mathfrak{A}$, is a model of Σ, the proof of the theorem is completed.

Using certain results from set theory about the exponentiation of infinite cardinals various consequences can be deduced from these theorems. We single out just one general result for mention here. Not surprisingly, since it involves cardinal exponentiation, it depends on the generalized continuum hypothesis.

THEOREM 5.4. (G.C.H.) *If Σ admits $\langle\alpha,\beta\rangle$, and $\langle\gamma,\delta\rangle$ are cardinals such that*

$$\delta\geqslant\operatorname{card}(\Sigma) \quad \textit{and} \quad \alpha\geqslant\gamma\geqslant\delta\geqslant\beta,$$

then Σ admits $\langle\gamma,\delta\rangle$.

PROOF. If $\delta=\beta$, this is an immediate consequence of theorem 5.1. Suppose then that $\beta<\delta$, then by the G.C.H. $\beta^{<\delta}=\delta$ and so by theorem 5.3, Σ

admits $\langle \alpha^{<\delta}, \delta \rangle$. But $\delta \leqslant \gamma \leqslant \alpha^{<\delta}$ and so, using theorem 5.1 again, we can conclude that Σ admits $\langle \gamma, \delta \rangle$.

We devote the next two sections to some more difficult two-cardinal theorems. Meanwhile we give an application of the simplest of the two cardinal theorems we have just proved to extend theorem 9.1.18 to infinite dimensional vector spaces. We will call such a vector space *normal*, if its dimension is equal to the cardinal of the field over which it is defined.

LEMMA 5.5. *Each infinite dimensional vector space over an infinite field is elementarily equivalent to a normal vector space of the same cardinality.*

PROOF. Let $\mathfrak{A}$ be an infinite dimensional vector space which is not normal. Suppose $\mathfrak{A}$ is of dimension α and $F_{\mathfrak{A}}$ is of cardinal β. (We are adopting here the notation for vector spaces that we used in ch. 9.) There are two cases to consider.

(i). $\beta < \alpha$. Let Σ be the set of all sentences of L true in $\mathfrak{A}$. With respect to the unary predicate letter F of L, Σ admits $\langle \alpha, \beta \rangle$ and so, by theorem 5.1, Σ admits $\langle \alpha, \alpha \rangle$. Notice that we can choose the model $\mathfrak{B}$ of Σ of cardinal α, with $\text{card}(F_{\mathfrak{B}}) = \alpha$, to be an elementary extension of $\mathfrak{A}$. $\mathfrak{B}$ is a normal vector space elementarily equivalent to $\mathfrak{A}$ and of the same cardinal as $\mathfrak{A}$.

(ii). $\alpha < \beta$. Let B be a basis for $\mathfrak{A}$ and let L′ be the language obtained from L by adding a unary predicate letter W. If Σ' is the set of all sentences of L′ true in $(\mathfrak{A}, B)$, Σ' admits $\langle \beta, \alpha \rangle$ and hence has a model $(\mathfrak{B}, B')$ of cardinal β with $\text{card}(B') = \beta$. Since $(\mathfrak{B}, B') \equiv (\mathfrak{A}, B)$, B' is a set of linearly independent vectors in $\mathfrak{B}$. Therefore $\mathfrak{B}$ is a vector space of cardinal β and dimension β elementary equivalent to $\mathfrak{A}$.

THEOREM 5.6. Let Δ *be any complete theory of infinite fields. Then the theory of vector spaces either of some specified finite dimension or of infinite dimension over fields which are models of* Δ *is complete.*

PROOF. For vector spaces of specified finite dimension the result has already been proved as theorem 9.1.8.

Suppose Σ is the theory of infinite dimensional vector spaces over fields which are models of Δ. Let $\mathfrak{A}, \mathfrak{B}$ be two models of Σ of cardinal α. By the previous lemma we can find normal vector spaces $\mathfrak{A}', \mathfrak{B}'$ of cardinal α, elementarily equivalent to $\mathfrak{A}$ and $\mathfrak{B}$ respectively. Since Δ is complete the fields

$$\mathfrak{A}'' = \langle F_{\mathfrak{A}'}, S_{\mathfrak{A}'}, P_{\mathfrak{A}'}, 0_{\mathfrak{A}'} \rangle,$$
$$\mathfrak{B}'' = \langle F_{\mathfrak{B}'}, S_{\mathfrak{B}'}, P_{\mathfrak{B}'}, 0_{\mathfrak{B}'} \rangle$$

are elementarily equivalent. By Kochen's ultralimit theorem they have

isomorphic ultralimits say $\mathfrak{A}''_\omega, \mathfrak{B}''_\omega$. Let $\mathfrak{A}'_\omega, \mathfrak{B}'_\omega$ be the ultralimits of $\mathfrak{A}', \mathfrak{B}'$ respectively, with respect to the same ultrafilter sequences. Since $\mathfrak{A}', \mathfrak{B}'$ are normal, $\mathfrak{A}'_\omega, \mathfrak{B}'_\omega$ are vector spaces of the same dimension over isomorphic fields. They are therefore isomorphic. We can immediately deduce from this that $\mathfrak{A} \equiv \mathfrak{B}$.

Since any two models of Σ of the same cardinal are elementarily equivalent, it follows from the Löwenheim–Skolem theorem that Σ is complete.

§ 6. Vaught's two-cardinal theorem

Our aim is to show that if Σ is a countable set of sentences which admits $\langle \alpha, \beta \rangle$ with $\aleph_0 \leqslant \beta < \alpha$, then Σ admits $\langle \aleph_1, \aleph_0 \rangle$. This result, due to Vaught, is not only of considerable interest by itself but also has important consequences for the theory of certain generalized quantifiers, which we explain in the next chapter.

Let L be some countable language and let L′ be the language obtained from L by adding the new unary predicate letter W. We associate with each formula ϕ of L a formula $\phi^{(W)}$ of L′, called the *relativization* of ϕ to W. $\phi^{(W)}$ is given by the following recursive definition:

(1). if ϕ is atomic, $\phi^{(W)} = \phi$.
(2). $(\neg\phi)^{(W)} = \neg\phi^{(W)}$.
(3). $(\phi \wedge \psi)^{(W)} = \phi^{(W)} \wedge \psi^{(W)}$.
(4). $((\exists v_n)\phi)^{(W)} = (\exists v_n)(W(v_n) \wedge \phi^{(W)})$.

EXERCISE 6.1. Show that

$$\vdash ((\forall v_n)\phi)^{(W)} \leftrightarrow (\forall v_n)(W(v_n) \rightarrow \phi^{(W)}).$$

It can be seen that $\phi^{(W)}$ is the formula which 'says the same thing' about the elements satisfying W as ϕ says about all elements.

Suppose now that U is a unary predicate letter, and that Σ is a set of sentences of L. For each formula ϕ of L, whose free variables are among $v_0, \ldots, v_{n-1}$, let ϕ_W be the sentence

$$(\forall v_0)\ldots(\forall v_{n-1})(W(v_0) \wedge \cdots \wedge W(v_{n-1}) \rightarrow [\phi \leftrightarrow \phi^{(W)}]).$$

Let Σ_W be the set of sentences of L′ consisting of the sentences in Σ, the sentences

$$(\forall v_0)[U(v_0) \rightarrow W(v_0)],$$
$$(\exists v_0)\neg W(v_0),$$

and all the sentences ϕ_W, where ϕ is a formula of L.

Vaught's two-cardinal theorem is a consequence of the next two lemmas, the first of which is straightforward.

LEMMA 6.2. *If Σ admits $\langle\alpha,\beta\rangle$ with $\aleph_0 \leqslant \beta < \alpha$, then Σ_W is consistent.*

PROOF. Let $\mathfrak{A}$ be a model of Σ of cardinal α in which $U_{\mathfrak{A}}$, the value of U, is a set of cardinal β. By theorem 4.3.1 there is an elementary substructure $\mathfrak{B}$ of $\mathfrak{A}$ of cardinal β such that

$$\mathfrak{A} \restriction U_{\mathfrak{A}} \subseteq \mathfrak{B} \prec \mathfrak{A}.$$

Let B be the domain of $\mathfrak{B}$. We claim that $\mathfrak{A}' = (\mathfrak{A}, B)$ is a model of Σ_W. Certainly $\mathfrak{A}'$ is a model of Σ. Since $U_{\mathfrak{A}} \subseteq B$, $\mathfrak{A}'$ is a model of

$$(\forall v_0)[U(v_0) \rightarrow W(v_0)],$$

and because $\operatorname{card}(B) < \operatorname{card}(A)$, $\mathfrak{A}'$ is a model of $(\exists v_0) \neg W(v_0)$.

The sentences

$$\{\phi_W : \phi \text{ is a formula of L}\}$$

say that the elements satisfying W form an elementary substructure, with respect to the language L, of the whole model, so, because $\mathfrak{B} \prec \mathfrak{A}$, $\mathfrak{A}'$ is also a model of all these sentences. Thus $\mathfrak{A}'$ is a model of Σ_W which is therefore consistent.

The next lemma is more complicated and its proof relies heavily on the results of the previous chapter.

LEMMA 6.3. *If Σ_W is consistent, then Σ admits $\langle\aleph_1,\aleph_0\rangle$.*

PROOF. Suppose Σ_W is consistent, then Σ_W has a model. It is clear that any model of Σ_W is infinite. Therefore if α is some uncountable limit beth number, it follows from corollary 11.4.4 that Σ_W has a special model of cardinal α, say

$$(\mathfrak{A}, W_{\mathfrak{A}}).$$

Since $(\mathfrak{A}, W_{\mathfrak{A}})$ is a model of Σ_W,

$$\mathfrak{A} \restriction W_{\mathfrak{A}} \prec \mathfrak{A}.$$

By the obvious extension of lemma 11.1.5, $\mathfrak{A}$ and $\mathfrak{A} \restriction W_{\mathfrak{A}}$ have the same cardinality and by the remark following theorem 11.4.5 they are both special models, and so by theorem 11.4.5, they are isomorphic. Let $R_{\mathfrak{A}}$ be a binary relation on $\mathfrak{A}$ which is an isomorphism between $\mathfrak{A}$ and $\mathfrak{A} \restriction W_{\mathfrak{A}}$. Let L'' be the language obtained from L' by adding a new binary predicate letter P, and let Δ be the set of all sentences of L'' which are true in

$$(\mathfrak{A}, W_{\mathfrak{A}}, R_{\mathfrak{A}}).$$

Δ has an infinite model, and so, by the downward Löwenheim–Skolem theorem, Δ has a countable model. By corollary 11.3.6 we can choose this countable model of Δ, say

$$(\mathfrak{B}, W_{\mathfrak{B}}, R_{\mathfrak{B}}),$$

so that $\mathfrak{B}$ is homogeneous.

Since $(\mathfrak{B}, W_{\mathfrak{B}}, R_{\mathfrak{B}})$ is a model of Δ and $\Sigma_W \subseteq \Delta$, we can conclude that $R_{\mathfrak{B}}$ is an isomorphism between $\mathfrak{B}$ and $\mathfrak{B} \restriction W_{\mathfrak{B}}$ which is a proper elementary substructure of $\mathfrak{B}$. Also $U_{\mathfrak{B}} \subseteq W_{\mathfrak{B}}$. Thus if $\mathfrak{B}_0 = \mathfrak{B} \restriction W_{\mathfrak{B}}$ and $\mathfrak{B}_1 = \mathfrak{B}$ we have

(1). $\mathfrak{B}_0$ is a proper elementary submodel of $\mathfrak{B}_1$.

(2). $\mathfrak{B}_0$ is isomorphic to $\mathfrak{B}_1$ and they are both homogeneous.

(3). $U_{\mathfrak{B}_0} = U_{\mathfrak{B}_1}$.

We now define an elementary chain $\langle \mathfrak{B}_\xi : \xi < \omega_1 \rangle$ of homogeneous, countable relational structures as follows:

$\mathfrak{B}_0$ has already been defined.

If $\xi = \zeta + 1$ is a successor ordinal with $\xi < \omega_1$, then $\mathfrak{B}_\xi$ is the relational structure related to $\mathfrak{B}_\zeta$ in exactly the same way that $\mathfrak{B}_1$ is related to $\mathfrak{B}_0$.

If $\delta < \omega_1$ is a limit ordinal then

$$\mathfrak{B}_\delta = \bigcup_{\xi < \delta} \mathfrak{B}_\xi .$$

$\mathfrak{B}_\delta$ is countable and the union of a chain of homogeneous structures and thus, by theorem 11.3.7 is homogeneous. Also each finite substructure of $\mathfrak{B}_\delta$ is isomorphic to a finite substructure of some $\mathfrak{B}_\xi$ with $\xi < \delta$, and hence to a finite substructure of $\mathfrak{B}_0$. We can therefore conclude from the isomorphism theorem that $\mathfrak{B}_\delta$ is isomorphic to $\mathfrak{B}_0$.

This completes our definition of the chain $\langle \mathfrak{B}_\xi : \xi < \omega_1 \rangle$. Now let

$$\mathfrak{B}_{\omega_1} = \bigcup_{\xi < \omega_1} \mathfrak{B}_\xi .$$

Then

$$U_{\mathfrak{B}_0} = U_{\mathfrak{B}_{\omega_1}}$$

and hence $U_{\mathfrak{B}_{\omega_1}}$ is of cardinal $\aleph_0$, while $\mathfrak{B}_{\omega_1}$ is clearly of cardinal $\aleph_1$. But $\mathfrak{B}_0 \prec \mathfrak{B}_{\omega_1}$ and so $\mathfrak{B}_{\omega_1}$ is a model of Σ. Hence we have proved that Σ admits $\langle \aleph_1, \aleph_0 \rangle$.

From these two lemmas we can immediately deduce

THEOREM 6.4. VAUGHT'S TWO-CARDINAL THEOREM

If Σ is a countable set of sentences containing the unary predicate letter U, which admits the pair of cardinals $\langle \alpha, \beta \rangle$ with $\aleph_0 \leqslant \beta < \alpha$, then Σ admits $\langle \aleph_1, \aleph_0 \rangle$.

The reason for splitting the proof of this theorem up in the way we have done will become apparent in the next chapter.

In a similar way we can deduce

THEOREM 6.5. *Let $\aleph_\alpha$ be a regular cardinal and let Σ be a countable set of sentences containing the binary predicate letter P which has an $\aleph_\alpha$-like model. Then Σ has an $\aleph_1$-like model.*

PROOF. By theorem 4.1 we need only consider the case $\alpha > 0$. Let $\mathfrak{A}_1$ be an $\aleph_\alpha$-like model of Σ, and for simplicity suppose that $<$ is the value of P in $\mathfrak{A}_1$.

By theorem 4.3.1, $\mathfrak{A}_1$ has a countable elementary substructure, say $\mathfrak{B}_0$, with domain B_0. Since $\aleph_\alpha$ is regular, there is a proper initial segment, say X_0, of A_1 which contains B_0. Since $<$ is $\aleph_\alpha$-like, $\operatorname{card}(X_0) < \aleph_\alpha$, and so, using theorem 4.3.1 again, $\mathfrak{A}_1$ has an elementary substructure $\mathfrak{B}_1$ of cardinal $< \aleph_\alpha$ which extends $\mathfrak{A}_1 \restriction X_0$. Proceeding in this way, we construct an elementary chain $\langle \mathfrak{B}_n : n < \omega \rangle$ of elementary substructures of $\mathfrak{A}_1$ each of cardinal $< \aleph_\alpha$ such that the domain of $\mathfrak{B}_{n+1}$ contains an initial segment of A_1 which includes the domain of $\mathfrak{B}_n$.

Let $\mathfrak{A}_0 = \bigcup_{n \in \omega} \mathfrak{B}_n$. It is clear from the previous remarks that $\mathfrak{A}_0$ is a proper elementary substructure of $\mathfrak{A}_1$ of cardinal $< \aleph_\alpha$ whose domain is an initial segment of the domain of $\mathfrak{A}_1$.

Now let L′ be the language obtained from L by adding the unary predicate letter W. Let

$$\Sigma'_W = \Sigma \cup \{\phi_W : \phi \text{ is a formula of L}\}$$
$$\cup \{(\forall v_0)[W(v_0) \to (\forall v_1)[P(v_1, v_0) \to W(v_1)]], (\exists v_0) \neg W(v_0)\}.$$

Then if A_0 is the domain of $\mathfrak{A}_0$,

$$\mathfrak{A}'_1 = (\mathfrak{A}_1, A_0)$$

can be seen to be a model of Σ'_W and so Σ'_W is consistent.

We now proceed as in the proof of lemma 6.3 and conclude that we can find isomorphic, countable, homogeneous models $\mathfrak{B}'_0, \mathfrak{B}'_1$ of Σ'_W such that $\mathfrak{B}'_0 \prec \mathfrak{B}'_1$ and the domain of $\mathfrak{B}'_0$ is an initial segment of the domain of $\mathfrak{B}'_1$ with respect to the ordering given by P. We can thus construct an elementary chain $\langle \mathfrak{B}'_\xi : \xi < \omega_1 \rangle$ of structures as in the proof of lemma 6.3, and we can conclude that

$$\mathfrak{B}_{\omega_1} = \bigcup_{\xi < \omega_1} \mathfrak{B}'_\xi$$

is an $\aleph_1$-like model of Σ.

§7. Chang's two-cardinal theorem

In this section we discuss C. C. Chang's generalization of Vaught's two-cardinal theorem to larger cardinals. The basic idea of the proof is the same as that used in the proof of Vaught's theorem. At one point in this proof we used the fact that the union of a countable chain of countable homogeneous structures is itself homogeneous. This does not extend to larger cardinals, and to get over this difficulty Chang had to resort to an ingenious device which we describe here.

We shall prove Chang's result only for countable languages. Its extension to uncountable languages is easy and we state this extension as a corollary to the main theorem. So we assume that L is a countable language with a unary predicate letter U. Let L′ be the language obtained from L by adding the binary predicate P and let L″ be the language obtained from L′ by adding the unary predicate letter W.

Let Σ be a set of sentences of L. Σ_P is the set of sentences of L′ consisting of all the sentences of Σ, the sentence

$$\sigma = (\forall v_0)(\forall v_1)[P(v_0,v_1) \to U(v_0) \wedge U(v_1)],$$

and for each formula ϕ of L the sentence ϕ_P which is the closure of

$$U(z_1) \wedge \cdots \wedge U(z_m) \wedge \phi(z_1) \wedge \cdots \wedge \phi(z_m) \\ \to (\exists v_k)[P(v_k,z_1) \wedge \cdots \wedge P(v_k,z_m) \wedge (\forall v_{k+1})[P(v_k,v_{k+1}) \to \phi(v_{k+1})]]$$

where none of the variables $z_1,\ldots,z_m,v_k,v_{k+1}$ occur in ϕ.

Then we let Σ'' be the set of sentences of L″ consisting of all the sentences of Σ_P, the sentences

$$(\forall v_0)[U(v_0) \to W(v_0)],$$
$$(\exists v_0) \neg W(v_0),$$

and for each formula ϕ of L′ whose free variables are among $v_0,\ldots,v_{n-1}$, the sentence

$$\phi_W = (\forall v_0)\ldots(\forall v_{n-1})[W(v_0) \wedge \cdots \wedge W(v_{n-1}) \to [\phi \leftrightarrow \phi^{(W)}]].$$

Thus Σ'' is the set of sentences obtained from Σ_P in exactly the same way that we obtained Σ_W from Σ in the previous section.

To point the similarity with Vaught's theorem we present the proof of Chang's theorem as an immediate consequence of two preliminary lemmas.

LEMMA 7.1. *If Σ admits the pair of cardinals $\langle \alpha,\beta \rangle$ with $\aleph_0 \leqslant \beta < \alpha$, then Σ'' is consistent.*

PROOF. Suppose Σ admits $\langle\alpha,\beta\rangle$ and that $\mathfrak{A}$ is a model of Σ of cardinal α, with $\text{card}(U_{\mathfrak{A}})=\beta$.

Since $U_{\mathfrak{A}}$ is infinite, there is a binary relation $P_{\mathfrak{A}}$ on $U_{\mathfrak{A}}$ which indexes all the finite subsets of $U_{\mathfrak{A}}$ in the sense that for each finite subset $\{u_1,\ldots,u_m\}$ of $U_{\mathfrak{A}}$, there is some $u\in U_{\mathfrak{A}}$ such that for all $a\in A$,

$$\langle u,a\rangle\in P_{\mathfrak{A}} \quad \text{iff} \quad a\in\{u_1,\ldots,u_m\}.$$

It is then clear that

$$\mathfrak{A}'=(\mathfrak{A},P_{\mathfrak{A}})$$

is a model of Σ_P.

The remainder of the proof of this lemma resembles the proof of lemma 6.2.

By theorem 4.3.1 there is an elementary substructure $\mathfrak{B}$ of $\mathfrak{A}$ of cardinal β such that

$$\mathfrak{A}\restriction U_{\mathfrak{A}}\subseteq\mathfrak{B}\prec\mathfrak{A}.$$

Let B be the domain of $\mathfrak{B}$. Then

$$\mathfrak{A}''=(\mathfrak{A}',B)$$

is a model of Σ'', which is therefore consistent.

Once again, the difficult part of the proof is contained in the second lemma. In this case this lemma depends on the generalized continuum hypothesis.

LEMMA 7.2. (G.C.H.) *If Σ'' is consistent then for any uncountable regular cardinal δ, Σ admits $\langle\delta^+,\delta\rangle$.*

PROOF. If Σ'' is consistent, Σ'' has a model. All models of Σ'' must be infinite. Therefore by theorem 11.3.4, if δ is an uncountable regular cardinal, Σ'' has a δ-saturated model of cardinal δ, say

$$\mathfrak{A}_0''=(\mathfrak{A}_0,R_0,W_0),$$

where $\mathfrak{A}_0$ is a realization of L. Since $\mathfrak{A}_0''$ is a model of Σ'',

$$\mathfrak{B}_0'=\mathfrak{A}_0'\restriction W_0\prec\mathfrak{A}_0',$$

where $\mathfrak{A}_0'=(\mathfrak{A}_0,R_0)$.

From this it is clear that W_0 is infinite and so by lemma 11.1.5 is of cardinal δ. Similarly U_0, the value of U in $\mathfrak{A}_0$ is of cardinal δ. By theorem 11.1.9 and the remarks preceding it, since $\mathfrak{A}_0''$ is δ-saturated, so too are $\mathfrak{A}_0'$ and $\mathfrak{B}_0'$, and therefore, because they are both of cardinal δ, by theorem 11.3.1 they are isomorphic.

Notice also that since the sentences

$$(\forall v_0)[U(v_0)\rightarrow W(v_0)]$$

and

$$(\forall v_0)(\forall v_1)[P(v_0,v_1)\rightarrow U(v_0)\wedge U(v_1)]$$

are in Σ'', the predicate letters U and P have the same values in $\mathfrak{B}'_0$ as in $\mathfrak{A}'_0$.

We now set about defining an elementary chain $\langle \mathfrak{B}'_\xi : \xi<\delta^+\rangle$ of relational structures such that for all $\nu<\delta^+$:

For each $\xi<\nu$, $\mathfrak{B}'_\xi$ is isomorphic to $\mathfrak{B}'_0$ and whenever $\eta<\xi$, $\mathfrak{B}'_\xi$ is a proper elementary extension of $\mathfrak{B}'_\eta$ in which U and P have the same values as in $\mathfrak{B}'_0$. [*]

$\mathfrak{B}'_0$ has already been defined.

If $\eta=\zeta+1$ is a successor ordinal with $\eta<\delta^+$ and for $\xi<\eta$, $\mathfrak{B}'_\xi$ has already been defined so as to satisfy [*], we let $\mathfrak{B}'_\eta$ be the relational structure which is related to $\mathfrak{B}'_\zeta$ in exactly the same way that $\mathfrak{B}'_1$ is related to $\mathfrak{B}'_0$.

We now come to the tricky part of the proof; namely the recursion step for limit ordinals. Suppose μ is a limit ordinal $\mu<\delta^+$ and that for $\xi<\mu$ $\mathfrak{B}'_\xi$ has been defined so as to satisfy [*]. Let

$$\mathfrak{B}'=\bigcup_{\xi<\mu}\mathfrak{B}'_\xi.$$

Clearly $\mathfrak{B}'$ is an elementary extension of each $\mathfrak{B}'_\xi$, with $\xi<\mu$, and is of cardinal δ. Further the value of U in $\mathfrak{B}'$ is the same as in $\mathfrak{B}'_0$, namely U_0. We are going to show that there is an elementary embedding of $\mathfrak{B}'$ into $\mathfrak{B}'_0$ which maps U_0 onto U_0.

Let $u\in {U_0}^\delta$ and $a\in {B'}^\delta$ be enumerations of U_0 and B' respectively, where B' is the domain of $\mathfrak{B}'$. We are going to define, by a second recursion, two sequences $b\in {B_0}^\delta$ and $b'\in {B'}^\delta$ such that for all $\nu<\delta$:

(i). $(\mathfrak{B}'_0,b\restriction \nu+1)\equiv(\mathfrak{B}',b'\restriction \nu+1)$, with respect to the extended language $L'_{\nu+1}$

and

(ii). $b'_\nu=a_{\eta+n}$, if $\nu=\eta+2n$ with $n<\omega$,

and

$b_\nu=u_{\eta+n}$, if $\nu=\eta+2n+1$ with $n<\omega$.

where in each case η is the largest limit ordinal $\leqslant\nu$.

Suppose $\lambda<\delta$, and the sequences b and b' have been defined for $\nu<\lambda$ so as to satisfy (i) and (ii). It follows that, with respect to the extended language L'_λ,

$$(\mathfrak{B}'_0,b\restriction\lambda)\equiv(\mathfrak{B}',b'\restriction\lambda). \tag{1}$$

There are two cases to consider. First suppose that $\lambda=\eta+2n$, with $n<\omega$. In this case we let

$$b'_\lambda=a_{\eta+n}.$$

Let Δ be the set of all formulas ϕ of F_λ such that

$$(\mathfrak{B}', b' \restriction \lambda) \models \phi[b'_\lambda].$$

Δ is finitely satisfiable in $(\mathfrak{B}', b' \restriction \lambda)$ and therefore, by (1), Δ is finitely satisfiable in $(\mathfrak{B}'_0, b \restriction \lambda)$. But $\mathfrak{B}'_0$ is δ-saturated and $\lambda < \delta$ and so Δ is satisfiable in $(\mathfrak{B}'_0, b \restriction \lambda)$. We let b_λ be any element of B_0 which satisfies Δ in $(\mathfrak{B}_0, b \restriction \lambda)$. It follows that

$$(\mathfrak{B}'_0, b \restriction \lambda+1) \equiv (\mathfrak{B}', b' \restriction \lambda+1),$$

with respect to the extended language $\mathrm{L}'_{\lambda+1}$ and so (i) and (ii) hold with ν replaced by λ.

Secondly, suppose $\lambda = \eta + 2n + 1$ with $n < \omega$. This case is more difficult. We let

$$b_\lambda = u_{\eta+n}$$

and we let Δ be the set of those formulas ϕ of F_λ such that

$$(\mathfrak{B}'_0, b \restriction \lambda) \models \phi[b_\lambda].$$

Δ is finitely satisfiable in $(\mathfrak{B}'_0, b \restriction \lambda)$, and so Δ is also finitely satisfiable in $(\mathfrak{B}', b' \restriction \lambda)$. We would like to be able to conclude that Δ is satisfiable in $(\mathfrak{B}', b' \restriction \lambda)$, but since we do not know that $\mathfrak{B}'$ is δ-saturated we cannot do this straight away. We have now reached the crux of the proof where Chang's device of introducing the extra predicate letter P to index finite subsets of the value of U plays its part.

Note at first that since $b_\lambda \in U_0$, the formula $U(v_0)$ is in Δ, and also that $\mathrm{card}(\Delta) < \delta$.

Let $\langle \Delta_\beta : \beta < \alpha \rangle$ be an enumeration of the finite subsets of Δ, and for each $\beta < \alpha$ let $\phi_\beta(v_0)$ be the conjunction of the formulas in Δ_β and the formula $U(v_0)$. Then for $\beta < \alpha$

$$(\mathfrak{B}', b' \restriction \lambda) \models (\exists v_0)\, \phi_\beta(v_0). \tag{2}$$

$\phi_\beta(v_0)$ is a formula of L'_λ and so contains a finite number of constants of L'_λ, say $c_{\zeta_1}, \ldots, c_{\zeta_k}$ with $\zeta_1, \ldots, \zeta_k < \lambda$. Thus from (2) it follows that

$$\mathfrak{B}' \models (\exists v_0)\, \phi_\beta(v_0)[b'_{\zeta_1}, \ldots, b'_{\zeta_k}].$$

There is some ordinal $\zeta_\beta < \mu$ such that

$$b'_{\zeta_1}, \ldots, b'_{\zeta_k} \in B'_{\zeta_\beta},$$

and so, because

$$\mathfrak{B}'_{\zeta_\beta} \prec \mathfrak{B}',$$

we have

$$\mathfrak{B}'_{\zeta_\beta} \models (\exists v_0)\, \phi_\beta(v_0)[b'_{\zeta_1},\ldots,b'_{\zeta_k}].$$

Hence there is some $u_\beta \in U_0$ such that

$$\mathfrak{B}' \models \phi_\beta[u_\beta, b'_{\zeta_1},\ldots,b'_{\zeta_k}].$$

Observe that if $\Delta_\beta \subseteq \Delta_\gamma$, then

$$\vdash \phi_\gamma(v_0) \to \phi_\beta(v_0),$$

and so, since $u_\gamma \in U_0 \subseteq B'_{\zeta_\beta}$,

$$\mathfrak{B}'_{\zeta_\beta} \models \phi_\beta[u_\gamma, b'_{\zeta_1},\ldots,b'_{\zeta_k}].$$

Now for each $\beta < \alpha$, let

$$\Gamma_\beta = \{P(v_0,d_\gamma) : \Delta_\beta \subseteq \Delta_\gamma\} \cup \{(\forall v_1)[P(v_0,v_1) \to \phi_\beta(v_1,c_{\zeta_1},\ldots,c_{\zeta_k})]\}.$$

Using the fact that $\mathfrak{B}'_{\zeta_\beta}$ is a model of Σ_P and hence of all the sentences ϕ_P, it can be seen that Γ_β is finitely satisfiable in $\mathfrak{B}'_{\zeta_\beta}$ when the constants d_γ are interpreted as referring to the elements u_γ, and $c_{\zeta_1},\ldots,c_{\zeta_k}$ are interpreted as referring to $b_{\zeta_1},\ldots,b_{\zeta_k}$. But $\mathfrak{B}'_{\zeta_\beta}$ is δ-saturated and so Γ_β is satisfiable in $\mathfrak{B}'_{\zeta_\beta}$, say by the element $w_\beta \in U_0$. Thus for each $\beta < \alpha$ we can find some $w_\beta \in U_0$ such that

$$\mathfrak{B}'_{\zeta_\beta} \models (\forall v_1)(P(v_0,v_1) \to \phi_\beta(v_1))[w_\beta, b_{\zeta_1},\ldots,b_{\zeta_k}] \tag{3}$$

and

$$\mathfrak{B}'_{\zeta_\beta} \models P[w_\beta, u_\gamma] \quad \text{whenever} \quad \Delta_\beta \subseteq \Delta_\gamma.$$

Let $\{e_\beta : \beta < \alpha\}$ be a new set of constant symbols and let

$$\Gamma = \{P(e_\beta, v_0) : \beta < \alpha\}.$$

Γ is a set of sentences of the language obtained from L′ by adding the constants $\{e_\beta : \beta < \alpha\}$.

Suppose $\{P(e_{\beta_i}, v_0) : 1 \leqslant i \leqslant m\}$ is a finite subset of Γ. Let $\gamma < \alpha$ be the ordinal such that

$$\Delta_\gamma = \Delta_{\beta_1} \cup \cdots \cup \Delta_{\beta_m}.$$

Then for $1 \leqslant i \leqslant m$,

$$\mathfrak{B}'_{\zeta_{\beta i}} \models P[w_{\beta_i}, u_\gamma].$$

Therefore, because

$$\mathfrak{B}'_0 \prec \mathfrak{B}'_{\zeta_{\beta i}},$$

and $w_{\beta_i}, u_\gamma \in U_0 \subseteq B'_0$, for $1 \leqslant i \leqslant m$,

$$\mathfrak{B}'_0 \models P[w_{\beta_i}, u_\gamma].$$

This shows that Γ is finitely satisfiable in

$$(\mathfrak{B}'_0,w),$$

where $w=\langle w_\beta:\beta<\alpha\rangle\in U_0{}^\alpha$. But $\mathfrak{B}'_0$ is δ-saturated and $\alpha<\delta$. Hence Γ is satisfiable in $(\mathfrak{B}'_0,w)$ and so there is some $u\in U_0$ such that for all $\beta<\alpha$,

$$\mathfrak{B}'_0\models P[w_\beta,u].$$

We show that u satisfies Δ in $(\mathfrak{B}',b'\restriction\lambda)$. Clearly it is sufficient to show that for all $\beta<\alpha$,

$$(\mathfrak{B}',b'\restriction\lambda)\models\phi_\beta[u].$$

So suppose $\beta<\alpha$. Then $\mathfrak{B}'_0\models P[w_\beta,u]$ and therefore, since $\mathfrak{B}'_0\prec\mathfrak{B}'_{\zeta_\beta}$

$$\mathfrak{B}'_{\zeta_\beta}\models P[w_\beta,u],$$

whence by (3),

$$\mathfrak{B}'_{\zeta_\beta}\models\phi_\beta[u,b_{\zeta_1},\ldots,b_{\zeta_k}],$$

and since $\mathfrak{B}'_{\zeta_\beta}\prec\mathfrak{B}'$,

$$\mathfrak{B}'\models\phi_\beta[u,b_{\zeta_1},\ldots,b_{\zeta_k}],$$

that is,

$$(\mathfrak{B}',b'\restriction\lambda)\models\phi_\beta[u].$$

Hence, at long last, we have shown that Δ is satisfiable in $(\mathfrak{B}',b'\restriction\lambda)$. So if we put

$$b'_\lambda=u,$$

(i) and (ii) will hold with ν replaced by λ.

Thus the sequences b and b' can be defined so as to satisfy (i) and (ii) for all $\nu<\delta$. It follows that

$$(\mathfrak{B}'_0,b)\equiv(\mathfrak{B}',b'),$$

with b' an enumeration of B' and with U_0 contained in the range of b. Thus the map $h:\mathfrak{B}'\to\mathfrak{B}'_0$ given by

$$h(b'_\nu)=b_\nu \quad \text{for} \quad \nu<\delta,$$

is an elementary embedding of $\mathfrak{B}'$ into $\mathfrak{B}'_0$ which maps U_0 onto U_0.

From this we can deduce that $\mathfrak{B}'$ has an elementary extension $\mathfrak{B}'_\mu$ which is isomorphic to $\mathfrak{B}'_0$. This completes the recursion step for limit ordinals in our definition of the elementary chain $\langle\mathfrak{B}'_\xi:\xi<\delta^+\rangle$. It follows that we can define this chain so as to satisfy [*] for all $\xi<\delta^+$.

Let

$$\mathfrak{A}'=(\mathfrak{A},R_{\mathfrak{A}})=\bigcup_{\xi<\delta^+}\mathfrak{B}'_\xi.$$

Clearly $\mathfrak{A}$, being the union of a properly increasing chain of δ^+ structures each of cardinal δ, is of cardinal δ^+. The value of U in $\mathfrak{A}$ is U_0 and so is a

set of cardinal δ. $\mathfrak{A}'$ is an elementary extension of each $\mathfrak{B}'_\xi$ and so is a model of Σ_P. Therefore $\mathfrak{A}$ is a model of Σ.

This completes the proof that Σ admits $\langle\delta^+,\delta\rangle$.

THEOREM 7.3. (G.C.H.) CHANG'S TWO-CARDINAL THEOREM
If Σ is a countable set of sentences which admits the pair of cardinals $\langle\alpha,\beta\rangle$ with $\aleph_0 \leqslant \beta < \alpha$, then for each regular cardinal δ, Σ admits $\langle\delta^+,\delta\rangle$.

PROOF. In the case $\delta = \aleph_0$ this is Vaught's two-cardinal theorem, and in the case $\delta > \aleph_0$ this is an immediate consequence of the two preceding lemmas.

The extension of the proof of theorem 7.3 to uncountable languages is straightforward and we have

LEMMA 7.4. (G.C.H.) *If Σ is a set of sentences which admits the pair of cardinals $\langle\alpha,\beta\rangle$ with $\aleph_0 \leqslant \beta < \alpha$, then for any regular cardinal $\delta > card(\Sigma)$, Σ admits $\langle\delta^+,\delta\rangle$.*

PROOF. The proof of this lemma is immediate from the proof of the preceding theorem.

§ 8. Historical and bibliographical remarks

Gödel's proof of the completeness of predicate calculus can be found in GÖDEL [1930]. The suggestion that his proof could be viewed as an ultraproduct construction was made to us by J. N. Crossley. Herbrand's theorem can be found in HERBRAND [1930]. The syntactic proof given there contains errors but it has recently been corrected (see DREBEN, ANDREWS and AANDERAA [1963], DREBEN and AANDERAA [1964], and DREBEN and DENTON [1966]).

Skolem's construction of a countable non-standard model of arithmetic was first presented in SKOLEM [1934] and again in SKOLEM [1965]. The relation between Skolem's construction and ultraproducts is discussed by SCOTT [1961].

Theorem 3.3 is due to MACDOWELL and SPECKER [1961] and its application to theorem 4.1 was pointed out by FUHRKEN [1964]. The two-cardinal theorems of §5 are due to CHANG and KEISLER [1962]. Vaught's two-cardinal theorem is proved in MORLEY and VAUGHT [1962]. Our version of the proof is based on a remark by FUHRKEN [1964] p. 293, as also is its use to prove theorem 6.5. Chang's two-cardinal theorem is proved in CHANG [1965].

CHAPTER 13

GENERALIZED QUANTIFIERS

In everything we have done so far the underlying languages have been based on first order predicate calculus. We have seen that these languages have several very convenient properties. We can sum these up as follows.

1. The completeness theorem. The importance of this result is not so much that some particular finite set of axioms is sufficient for proving all universally valid sentences, but that this shows that the universally valid sentences form a recursively enumerable set. We used this fact when we showed that recursively axiomatizable, complete theories are decidable.

2. The compactness theorem. The reader will already have come across several important applications of this theorem. It will be evident that if this theorem broke down completely the construction of models would be very awkward.

3. Löwenheim–Skolem theorems. The usefulness of these theorems should also be apparent. In particular it should be noted that Vaught's test for completeness depends on them.

4. Characterizability of many mathematical notions. There would not be much point in pursuing the study of first order predicate logic, if the languages we were investigating could not be used to say anything interesting. But, as we have seen, many algebraic notions can be described by sentences of first order predicate logic.

Despite all these 'nice' properties, first order predicate calculus does not quite fulfil the function of a universal language or *characteristica universalis* for all of mathematics. For example, as we have already noted, only finite structures can be described categorically by sets of sentences of first order predicate calculus, and certain notions, such as that of a well-ordering cannot be described at all.

It is natural, therefore, to investigate the extent to which classical predicate calculus can be extended in ways which preserve as many of the 'nice' properties as possible while increasing its expressive power. We describe two directions such investigations have taken in the last two chapters of this

book. But it is not possible, in the space available, to discuss even these in great detail, and we will often leave it to the reader to pursue his study in the original publications of research in this area.

§ 1. Generalized quantifiers

The first thing to notice is that some of the 'defects' of predicate calculus are immediate consequences of its 'nice' properties. Thus we have seen that the compactness theorem implies the upward Löwenheim–Skolem theorem and that this implies that infinite structures cannot be characterized by categorical axioms. So at once we have to make a choice about which properties we want to preserve. We now set about describing some of the choices that can be made.

In this chapter we discuss the effect of adding new quantifiers to the language L. The possibility of doing this was originally suggested by MOSTOWSKI [1957].

We can regard quantifiers as operators which pick out certain subsets from among all the subsets of the domain of a given realization. Thus the existential quantifier picks out the non-empty sets, and the universal quantifier just the whole domain itself.

Our first thought, therefore, might be to identify generalized quantifiers with functions which associate with each realization a collection of subsets of its domain. Let λ be such a function. We let L′ be the language obtained from L by dropping the existential quantifier and replacing it by the general quantifier symbol Q. Then we interpret Q as the generalized quantifier λ, if we replace the satisfaction clause for $\exists$, by the following clause for Q:

$$\mathfrak{A} \models_x (Qv_n)\,\phi \quad \text{iff} \quad \{a \in A : \mathfrak{A} \models_{x(n/a)} \phi\} \in \lambda(\mathfrak{A}).$$

A moment's reflection will however show that in order to meet our intuitive idea of what a quantifier is we must place some restrictions on the function λ.

First, since quantifiers operate on *sets* the values of λ should be independent of the structure $\mathfrak{A}$ and should depend only on the set A which is the domain of $\mathfrak{A}$. Secondly, quantifiers should not distinguish between elements but should pick out sets only according to their 'size' compared with that of the domain. Thus if π is a one–one map of A onto B, and $X \in \lambda(A)$ then we require that $\pi[X] \in \lambda(B)$.

These considerations lead us to give the following somewhat cumbersome definition of a *generalized quantifier* or *interpretation of a quantifier symbol.* An *interpretation of a quantifier symbol* Q is a map μ_Q defined on all ordered

triples $\langle \mathfrak{m},\mathfrak{n},\mathfrak{p}\rangle$ of cardinals, such that $\mathfrak{m}+\mathfrak{n}=\mathfrak{p}$ and taking the values 0 and 1. Corresponding to this interpretation we have the following satisfaction definition for the quantifier symbol Q:

$$\mathfrak{A} \vDash_x (Qv_n)\,\phi \quad \text{iff} \quad \mu_Q(\mathfrak{m},\mathfrak{n},\mathfrak{p})=1,$$

where

$$\mathfrak{m}=\mathrm{card}(\{a\in A:\mathfrak{A}\vDash_{x(n/a)}\phi\}),$$
$$\mathfrak{n}=\mathrm{card}(\{a\in A:\mathfrak{A}\vDash_{x(n/a)}\neg\phi\})$$

and

$$\mathfrak{p}=\mathrm{card}(A).$$

It can be seen that this definition corresponds with our intuitive idea and that it is the broadest definition which does so.

EXAMPLES

1.1. If

$$\mu_Q(\mathfrak{m},\mathfrak{n},\mathfrak{p})=1 \quad \text{iff} \quad \mathfrak{m}\geqslant 1,$$

then Q is the existential quantifier.

1.2. If

$$\mu_Q(\mathfrak{m},\mathfrak{n},\mathfrak{p})=1 \quad \text{iff} \quad \mathfrak{n}=0,$$

then Q is the universal quantifier.

1.3. If

$$\mu_Q(\mathfrak{m},\mathfrak{n},\mathfrak{p})=1 \quad \text{iff} \quad \mathfrak{m}=k, \quad \text{with } k\in\omega,$$

then Q is the quantifier $\exists!^k$ which says 'there are exactly k'. Clearly this quantifier can be defined in terms of the existential quantifier.

We give further examples in the next section.

§ 2. The quantifiers Q_α

The quantifier Q_α is interpreted as saying 'there exist at least $\aleph_\alpha$' Q_α is thus the quantifier corresponding to the function μ_{Q_α} given by

$$\mu_{Q_\alpha}(\mathfrak{m},\mathfrak{n},\mathfrak{p})=1 \quad \text{iff} \quad \mathfrak{m}\geqslant\aleph_\alpha,$$

or, more straightforwardly, it is the quantifier defined by the satisfaction clause

$$\mathfrak{A}\vDash_x (Q_\alpha v_n)\phi \quad \text{iff} \quad \mathrm{card}(\{a\in A:\mathfrak{A}\vDash_{x(n/a)}\phi\})\geqslant\aleph_\alpha.$$

We let L_{Q_α} be the language obtained from L by adding the quantifier Q_α.

In a realization of L_{Q_α} of cardinal $<\aleph_\alpha$ the quantifier Q_α acts trivially and $(Q_\alpha x)\,\phi(x)$ is always false. To eliminate trivial complications, therefore, when we are talking about the language L_{Q_α} we assume that *all realizations are of cardinal* $\geqslant\aleph_\alpha$.

A simple argument shows that the compactness theorem, in its full generality, does not hold for any of the languages L_{Q_α}.

THEOREM 2.1. *For each* α, *there is a set* Σ *of* $\aleph_\alpha$ *sentences of* L_{Q_α}, *such that each finite subset of* Σ *has a model, but* Σ *itself does not have a model.*

PROOF. Let $\{P_\xi : \xi<\aleph_\alpha\}$ be a collection of $\aleph_\alpha$ unary predicate letters. We let

$$\begin{aligned}\Sigma = &\{(\exists v_0)[P_0(v_0)\wedge P_\xi(v_0)] : 0<\xi<\aleph_\alpha\}\\ &\cup\{\neg(\exists v_0)[P_\xi(v_0)\wedge P_\eta(v_0)] : 0<\xi<\eta<\aleph_\alpha\}\\ &\cup\{\neg(Q_\alpha v_0)\,P_0(v_0)\}.\end{aligned}$$

The first two sets of sentences making up Σ imply that the value of P_0 is a set of cardinal $\geqslant\aleph_\alpha$ contradicting $\neg(Q_\alpha v_0)P_0(v_0)$, thus Σ does not have a model. However, since only finitely many of the predicate letters P_ξ occur in any finite subset of Σ, it is easily seen that each finite subset of Σ has a model.

Although this shows that the full compactness theorem does not hold for any of the languages L_{Q_α} it still leaves open the possibility that it might hold for comparatively small sets of sentences. Let us say that a language is $\mathfrak{m}$-*compact* if whenever Σ is a set of sentences of the language of cardinal $\leqslant\mathfrak{m}$ such that each finite subset of Σ has a model, then Σ itself has a model. The previous theorem shows that L_{Q_α} is not $\aleph_\alpha$-compact; it leaves the possibility that L_{Q_α} is $\mathfrak{m}$-compact for some $\mathfrak{m}<\aleph_\alpha$. Not surprisingly, we can use the ultraproduct construction to obtain some results in this direction.

If $\mathfrak{n}$ is an infinite cardinal, we say that $\mathfrak{n}$ is *small for* $\aleph_\alpha$ if whenever $\{\mathfrak{m}_i : i\in I\}$ is a collection of cardinals each less than $\aleph_\alpha$ and $\operatorname{card}(I)\leqslant\mathfrak{n}$, we have

$$\prod_{i\in I}\mathfrak{m}_i<\aleph_\alpha.$$

THEOREM 2.2. ŁOŚ'S THEOREM FOR L_{Q_α}.
Let $\{\mathfrak{A}_i : i\in I\}$ *be a collection of relational structures, and let* F *be an ultrafilter on* I. *If* $\operatorname{card}(I)$ *is small for* $\aleph_\alpha$ *then for any formula* ϕ *of* L_{Q_α} *and any sequence* $f/F\in(\prod A_i/F)^\omega$

$$\prod\mathfrak{A}_i/F\vDash_{f/F}\phi \quad \textit{iff} \quad \{i\in I : \mathfrak{A}_i\vDash_{f(i)}\phi\}\in F.$$

PROOF. The proof is by induction on the number of logical symbols in ϕ. The proof for atomic formulas and of the induction steps for the connectives $\neg, \wedge$ and the quantifier $\exists$ are exactly the same as in the proof of Łoš's theorem for L (ch. 5, §2).

Suppose now that the theorem holds for the formula ψ. We will show that it holds also for the formula $\phi = (Q_\alpha v_n)\,\psi$.

For $i \in I$, let $B_i = \{a \in A_i : \mathfrak{A}_i \vDash_{f(i)(n/a)} \psi\}$. Clearly, by the induction hypothesis

$$\prod B_i/F = \{a \in \prod A_i/F : \prod \mathfrak{A}_i/F \vDash_{f/F(n/a)} \psi\}.$$

Let $X = \{i \in I : \mathfrak{A}_i \vDash_{f(i)} (Q_\alpha v_n)\,\psi\}$. If $X \in F$, then by theorem 6.2.2,

$$\operatorname{card}\Big(\prod_{i \in I} B_i/F\Big) = \operatorname{card}\Big(\prod_{i \in X} B_i/F \restriction X\Big).$$

By the definition of X, for $i \in X$, $\operatorname{card}(B_i) \geqslant \aleph_\alpha$ and hence $\operatorname{card}(\prod B_i/F \restriction X) \geqslant \aleph_\alpha$, and therefore $\operatorname{card}(\prod B_i/F) \geqslant \aleph_\alpha$. Thus

$$\prod \mathfrak{A}_i/F \vDash_{f/F} (Q_\alpha v_n)\,\psi\,.$$

Now suppose $X \notin F$. Then $I - X = Y \in F$ and

$$\operatorname{card}(\prod B_i/F) = \operatorname{card}(\prod B_i/F \restriction Y).$$

For $i \in Y$, $\operatorname{card}(B_i) < \aleph_\alpha$ and so

$$\operatorname{card}(\prod B_i/F) = \operatorname{card}(\prod B_i/F \restriction Y) \leqslant \prod_{i \in Y} \operatorname{card}(B_i) < \aleph_\alpha,$$

since $\operatorname{card}(I)$, and a fortiori $\operatorname{card}(Y)$, is small for $\aleph_\alpha$. Thus

$$\text{not} \quad \prod \mathfrak{A}_i/F \vDash_{f/F} (Q_\alpha v_n)\,\psi\,.$$

This completes the proof of the theorem.

THEOREM 2.3. *If* $\mathfrak{n}$ *is small for* $\aleph_\alpha$, L_{Q_α} *is* $\mathfrak{n}$*-compact.*

PROOF. We have just seen that if $\mathfrak{n}$ is small for $\aleph_\alpha$, Łoš's theorem applies to ultraproducts containing not more than $\mathfrak{n}$ factors. Therefore, using ultraproducts, we can prove theorem 2.3 in exactly the same way as we proved the compactness theorem for L (ch. 5, §4).

COROLLARY 2.4. *If* $\aleph_0$ *is small for* $\aleph_\alpha$, L_{Q_α} *is* $\aleph_0$*-compact.*

Clearly $\aleph_0$ is not small for any cardinal $\leqslant 2^{\aleph_0}$ and, in particular, $\aleph_0$ is not small for $\aleph_1$. Thus theorem 2.3 cannot be used to show that L_{Q_1} is $\aleph_0$-

compact. This result can however be proved by a method we indicate later in this chapter. The general question of which cardinals are small for $\aleph_\alpha$ is a problem in set theory and we do not discuss it here. The reader can find what is known about this question in BACHMANN [1967].

The downward Löwenheim–Skolem theorem for the languages L_{Q_α} is proved by a natural extension of the method that we used to prove the result for L.

THEOREM 2.5. *Let $\mathfrak{A}$ be a realization of L_{Q_α} of cardinal $\mathfrak{n}$, and let C be a subset of A of cardinal $\mathfrak{p}$. If $\mathfrak{m}$ is a cardinal $\geqslant$ the cardinal of the language L_{Q_α} and $\geqslant \aleph_\alpha$ such that $\mathfrak{p} \leqslant \mathfrak{m} \leqslant \mathfrak{n}$ then $\mathfrak{A} \restriction C$ has an extension $\mathfrak{B}$ of cardinal $\mathfrak{m}$ which is an elementary substructure of $\mathfrak{A}$.*

PROOF. Let $A^{<\omega}$ be the collection of all finite sequences of elements of A and let $\mathscr{F}$ be the set of all formulas of L_{Q_α}. We define a function

$$h: \mathscr{F} \times A^{<\omega} \to \mathscr{P}(A)$$

as follows:

Suppose $\phi(v_0,\ldots,v_n)$ is a formula of L_{Q_α} with free variables among $v_0,\ldots,v_n$ and $\boldsymbol{a} = \langle a_0,\ldots,a_{n-1}\rangle$ is a sequence of elements of A. There are three cases to consider:

(1). $\mathfrak{A} \models (Q_\alpha v_n)\, \phi[a_0,\ldots,a_{n-1}]$.

In this case the set $\{a \in A : \mathfrak{A} \models \phi[a_0,\ldots,a_{n-1},a]\}$ is of cardinal $\geqslant \aleph_\alpha$ and we let $h(\phi,\boldsymbol{a})$ be any subset of this set containing exactly $\aleph_\alpha$ elements.

(2). not $\mathfrak{A} \models (Q_\alpha v_n)\, \phi[a_0,\ldots,a_{n-1}]$ but $\mathfrak{A} \models (\exists v_n)\, \phi[a_0,\ldots,a_{n-1}]$.

In this case $\{a \in A : \mathfrak{A} \models \phi[a_0,\ldots,a_{n-1},a]\}$ is non-empty and we let $h(\phi,\boldsymbol{a})$ be any one-element subset of this set.

(3). not $\mathfrak{A} \models (\exists v_n)\, \phi[a_0,\ldots,a_{n-1}]$.

In this case we let $h(\phi,\boldsymbol{a})$ be the empty set.

Now we define the sequence $\langle B_n : n \in \omega\rangle$ of subsets of A as follows:

B_0 is any subset of A of cardinal $\mathfrak{m}$ which contains C.

$$B_{n+1} = \bigcup \{h(\phi,\boldsymbol{a}) : \phi \text{ is a formula of } L_{Q_\alpha} \text{ with free variables among } v_0,\ldots,v_n \text{ and } \boldsymbol{a} = \langle a_0,\ldots,a_{n-1}\rangle \text{ is a sequence of } n \text{ elements from } B_n.\}$$

Just as in the proof of the analogous result for L (theorem 4.3.1) it can be seen that for $n \in \omega$, $B_n \subseteq B_{n+1}$ and $\operatorname{card}(B_n) = \mathfrak{m}$. Therefore if $B = \bigcup_{n\in\omega} B_n$, $\operatorname{card}(B) = \mathfrak{m}$ and by construction $C \subseteq B$.

If we let $\mathfrak{B} = \mathfrak{A} \restriction B$, then an argument similar to that used in proving theorem 4.3.1 shows that $\mathfrak{B} \prec \mathfrak{A}$.

COROLLARY 2.6. THE DOWNWARD LÖWENHEIM–SKOLEM THEOREM FOR L_{Q_α}. *Let Σ be a set of sentences of L_{Q_α} of cardinal $\mathfrak{m}$. If Σ has a model of cardinal $\geqslant \mathfrak{m}+\aleph_\alpha$ then Σ has a model of cardinal $\mathfrak{m}+\aleph_\alpha$.*

Just as the case with L, we can use this result to prove an analogue of Vaught's test for L_{Q_α}.

THEOREM 2.7. VAUGHT'S TEST FOR L_{Q_α}.
Let Σ be a countable set of sentences of L_{Q_α} which is $\aleph_\alpha$-categorical, then Σ is complete.

PROOF. Note, that since we are adopting the convention that all realizations of L_{Q_α} must be of cardinal $\geqslant \aleph_\alpha$ we do not need to impose the restriction that Σ has only infinite models in the statement of this theorem. The proof is similar to that of the analogous result for L.

It should be noticed that 'complete' here means 'semantically complete' in the sense that all models of Σ are elementarily equivalent, that is complete according to our original definition of this concept. As yet we have not discussed the possibility of axiomatizing any of the languages L_{Q_α} and so we have no analogue of theorem 9.1.7 equating semantic and syntactic completeness. Indeed we can use theorem 2.7 to show that in one case it is impossible to find a recursive set of axioms for L_{Q_α}.

The quantifier Q_0 says 'there exist at least $\aleph_0$' or equivalently 'there exist infinitely many' and so $\neg Q_0$ says 'there are only finitely many'. The fact that we can in this way express the notion of a finite set in L_{Q_0} shows how much more powerful L_{Q_0} is than L. In fact L_{Q_0} contains a categorical set of axioms for arithmetic. Let *FA* be any usual set of axioms for first order formal arithmetic and let

$$FA^* = FA \cup \{(\forall v_0)\neg(Q_0 v_1)(v_1 < v_0)\},$$

then in any model of FA^* the elements form an ordered set of order type ω with respect to the natural ordering. Hence the model contains no non-standard elements and so is isomorphic to the standard model of arithmetic. This shows that FA^* is a categorical set of sentences and thus, by Vaught's test for L_{Q_0}, is semantically complete. We can deduce from this that there is no recursive axiomatization for L_{Q_0}:

THEOREM 2.8. *There is no recursive axiomatization of L_{Q_0} such that for any set Σ of sentences of L_{Q_0} those sentences provable from the assumptions Σ are precisely those sentences true in all models of Σ.*

PROOF. If there were such a recursive axiomatization the set of sentences true in all models of *FA** would form a recursively enumerable set, and hence the set Δ of those sentences of L_{Q_0} not containing Q_0 true in all models of *FA** would be a recursively enumerable set. Since *FA** is categorical Δ is the set of all sentences true in the standard model of arithmetic. But, as is well known, it follows from Gödel's incompleteness theorem for arithmetic that this set of sentences is not recursively enumerable. This contradiction proves the theorem.

Note that this last result leaves open the possibility that there is a recursive set of axioms for L_{Q_0} in the weaker sense that the axiomatization be sufficiently strong to prove all universally valid sentences but not necessarily strong enough to prove from Σ those sentences true in all models of Σ, for an arbitrary set Σ of sentences. To see that such a recursive axiomatization is not possible, it is sufficient to observe that by dropping the induction schema from *FA** we obtain a *finite* set of categorical axioms whose conjunction is σ, say. Then a sentence τ of L is true in the standard model of arithmetic iff $\sigma \to \tau$ is universally valid. Thus if the universally valid sentences of L_{Q_0} were recursively enumerable, so also would the sentences of L true in the standard model of arithmetic, and we again would have a contradiction.

Clearly theorem 2.8 can be generalized so as to show that there is no language L* such that

(a). there is an $\aleph_0$-categorical recursive set of axioms for arithmetic in L*

(b). every countable set of sentences of L* with an infinite model has a countable model,

and

(c). there is a recursive axiomatization for L*.

This is another restriction on the possible 'nice' extensions of L. In the next section we describe a language with a generalized quantifier for which (a) and (c) hold but (b) breaks down. Ordinary predicate calculus is a language with properties (b) and (c), and we have just seen that L_{Q_0} satisfies (a) and (b).

§ 3. The Chang quantifier

We let Q be the quantifier corresponding to the function μ_Q given by

$$\mu_Q(m,n,p) = 1 \quad \text{iff} \quad m = p,$$

or more straightforwardly, it is the quantifier defined by the satisfaction clause

$$\mathfrak{A} \vDash_x (Qv_n)\,\phi \quad \text{iff} \quad \text{card}(\{a \in A : \mathfrak{A} \vDash_{x(n/a)} \phi\}) = \text{card}(A).$$

Q is called the *Chang quantifier* after C. C. Chang, and sometimes the *equi-cardinal quantifier.*

Notice that in a finite structure Q behaves like the universal quantifier. To avoid trivial complications when talking about Q we shall adopt the convention that *all structures are infinite.* Also in a realization of cardinal $\aleph_\alpha$, Q behaves like the quantifier Q_α described in the previous section. It follows from this that many of the results about the quantifiers Q_α and the quantifier Q are intertranslatable although this is not true for any of the results mentioned in the previous section. It is however also possible to obtain a variant of Łoš's theorem for the language L_Q obtained from L by adding the quantifier Q. This we now do.

We say that the cardinal $\mathfrak{m}$ is $\mathfrak{n}$-*normal* if for any collection $\{\mathfrak{m}_\xi : \xi < \mathfrak{n}\}$ of cardinals each less than $\mathfrak{m}$, we have

$$\prod_{\xi < \mathfrak{n}} \mathfrak{m}_\xi < \mathfrak{m}^{\mathfrak{n}}.$$

THEOREM 3.1. ŁOŠ'S THEOREM FOR L_Q
Let $\{\mathfrak{A}_i : i \in I\}$ be a collection of relational structures each of cardinal $\mathfrak{m}$. If $\mathfrak{m}$ is card(I)-normal and F is a regular ultrafilter on I, then for any formula ϕ of L_Q and any sequence $f/F \in (\prod A_i/F)^\omega$

$$\prod \mathfrak{A}_i/F \vDash_{f/F} \phi \quad \textit{iff} \quad \{i \in I : \mathfrak{A}_i \vDash_{f(i)} \phi\} \in F.$$

PROOF. Notice that by theorem 6.3.21, since F is regular $\text{card}(\prod A_i/F) = \mathfrak{m}^{\mathfrak{n}}$.

With this observation in mind it is now easily seen that the theorem can be proved by adopting a method similar to that used to prove theorem 2.2.

THEOREM 3.2. *Let Σ be a set of sentences of L_Q each finite subset of which has a model of cardinal $\mathfrak{m}$. If $\mathfrak{m}$ is card(Σ)-normal then Σ has a model.*

PROOF. We can assume Σ is infinite. Let $I = S_\omega(\Sigma)$; since Σ is infinite $\text{card}(I) = \text{card}(\Sigma)$. By hypothesis for each $\Delta \in I$, there is a model $\mathfrak{A}_\Delta$ of Δ of cardinal $\mathfrak{m}$.

For $\Delta \in S_\omega(\Sigma)$, let

$$\Delta^* = \{\Delta' \in S_\omega(\Sigma) : \Delta \subseteq \Delta'\}.$$

It is easily seen that $\{\Delta^* : \Delta \in S_\omega(\Sigma)\}$ has the finite intersection property and

so can be extended to an ultrafilter F on I. We will show that F is a regular ultrafilter.

Let g be any one–one map of Σ onto $S_\omega(\Sigma)=I$. We define the map $f:I\to S_\omega(I)$ by

$$f(\Delta)=\{g(\sigma):\sigma\in\Delta\}.$$

Clearly f is one–one onto. Suppose $\Delta\in I$ and let

$$\Delta^+=\{\Delta'\in I:\Delta\in f(\Delta')\}.$$

Since f is onto, for some $\Delta_0\in I$, $\Delta\in f(\Delta_0)$. If $\Delta_0\subseteq\Delta'$,

$$f(\Delta_0)\subseteq f(\Delta')$$

and so $\Delta\in f(\Delta')$. Thus

$$\Delta_0{}^*=\{\Delta'\in I:\Delta_0\subseteq\Delta'\}\subseteq\{\Delta'\in I:\Delta\in f(\Delta')\}=\Delta^+$$

By construction $\Delta_0{}^*\in F$, therefore $\Delta^+\in F$. This shows that F is regular.

It follows from this that theorem 3.1 applies to the ultraproduct $\prod\mathfrak{A}_\Delta/F$. Suppose $\sigma\in\Sigma$, then $\{\sigma\}\in I$, say $\{\sigma\}=\Delta_0$. $\mathfrak{A}_{\Delta_0}\models\sigma$ and clearly if $\Delta_0\subseteq\Delta'$ then $\mathfrak{A}_{\Delta'}\models\sigma$. Hence

$$\Delta_0{}^*=\{\Delta'\in I:\Delta_0\subseteq\Delta'\}\subseteq\{\Delta'\in I:\mathfrak{A}_{\Delta'}\models\sigma\}.$$

Since $\Delta_0{}^*\in F$ it follows that

$$\{\Delta'\in I:\mathfrak{A}_{\Delta'}\models\sigma\}\in F$$

and therefore, by theorem 3.1, that $\prod\mathfrak{A}_\Delta/F\models\sigma$. Thus we have shown that $\prod\mathfrak{A}_\Delta/F$ is a model of Σ, which completes the proof of the theorem.

The usefulness of this last result will become apparent in the next section when we have described Fuhrken's technique for obtaining Löwenheim–Skolem theorems for the language L_Q. We will not discuss the general question of when a cardinal $\mathfrak{m}$ is $\mathfrak{n}$-normal; once again the reader can find information about this in BACHMANN [1967]. However we shall want to use the next result.

LEMMA 3.3. *For each infinite cardinal* $\mathfrak{n}$, $(2^{\mathfrak{n}})^+$ *is* $\mathfrak{n}$*-normal.*

PROOF. Suppose $\{\mathfrak{m}_\xi:\xi<\mathfrak{n}\}$ is a collection of cardinals each less than $(2^{\mathfrak{n}})^+$. Then for $\xi<\mathfrak{n}$, $\mathfrak{m}_\xi\leqslant 2^{\mathfrak{n}}$ and so

$$\prod_{\xi<\mathfrak{n}}\mathfrak{m}_\xi\leqslant(2^{\mathfrak{n}})^{\mathfrak{n}}=2^{\mathfrak{n}}<(2^{\mathfrak{n}})^+\leqslant((2^{\mathfrak{n}})^+)^{\mathfrak{n}}.$$

Note that it can be shown that $((2^{\mathfrak{n}})^+)^{\mathfrak{n}}=(2^{\mathfrak{n}})^+$.

§ 4. Fuhrken's reduction technique

In this section we describe a method, due to FUHRKEN [1964] for showing that questions about whether a set of sentences of L_Q has a model can be reduced to questions about whether an associated set of sentences not containing the Chang quantifier has a model of a special kind. The same techniques can be applied equally well to the languages L_{Q_κ}.

Suppose first that $\mathfrak{A}$ is a realization of L_Q of some successor cardinal $\mathfrak{m}^+$. Fuhrken's idea is the following: We pick some subset B of A of cardinal $\mathfrak{m}$ and for each definable subset X of A we choose a one-one map of X either into B or onto the whole of A. Since there are only $\mathfrak{m}^+$ definable subsets of A, we can index these functions by elements of A. In this way we define a ternary relation Y on A such that for each $a \in A$ the function g_a defined by

$$g_a(b)=c \quad \text{iff} \quad \langle a,b,c\rangle \in Y$$

is one of the functions mentioned above. Then instead of saying $(Qv_0)\,\phi(v_0)$, we say that there is some element which indexes a function mapping the elements satisfying ϕ onto the whole domain, and instead of $\neg(Qv_0)\,\phi(v_0)$ we say that there is some element which indexes a function mapping the elements satisfying ϕ into B. We now make these remarks more precise.

Let L be some first order language. We will assume L is countable; the extension to other cases will be obvious.

We let L_Q be the language obtained from L by adding the Chang quantifier Q, and we let L* be the language obtained from L by adding a new unary predicate letter U and a new ternary predicate letter F. We associate with each formula ϕ of L_Q a formula ϕ^* of L*, defined by the following recursion:

(1). if ϕ is atomic $\phi^*=\phi$,
(2). $(\neg\phi)^*=\neg\phi^*$,
(3). $(\phi\wedge\psi)^*=\phi^*\wedge\psi^*$,
(4). $((\exists v_n)\,\phi)^*=(\exists v_n)\,\phi^*$,
(5). $((Qv_n)\,\phi)^*=(\exists v_k)(\forall v_{k+1})(\exists v_n)[\phi^*\wedge F(v_k,v_n,v_{k+1})]$, where v_k is the first variable in the sequence $v_0,v_1,\ldots$ occurring after all those variables in ϕ^*.

Next we associate with each set Σ of sentences of L_Q a set Σ^* of sentences of L*. Σ^* consists of all the following sentences:

(a). All the sentences σ^* for $\sigma\in\Sigma$.

(b). The sentence which says that 'F is an indexed collection of one–one functions', namely

$$(\forall v_0)(\forall v_1)\{(\exists v_2)(\forall v_3)[F(v_0,v_1,v_3)\to v_3=v_2] \wedge (\exists v_2)(\forall v_3)[F(v_0,v_3,v_1)\to v_3=v_2]\}.$$

(c). For each formula of the form $(Qv_n)\,\psi$ which is a subformula of some sentence in Σ, the closure of

$$(\exists v_k)\{(\forall v_{k+1})(\exists v_n)[\psi^* \wedge F(v_k,v_n,v_{k+1})] \vee (\forall v_n)(\exists v_{k+1})[\psi^*\to U(v_{k+1})\wedge F(v_k,v_n,v_{k+1})]\},$$

where v_k is the first variable after v_n and all those in ψ^*.

This is the sentence which says that among the functions indexed by F there is one which either maps the elements satisfying ψ onto the whole domain or into the set of elements satisfying U.

The point of this translation is given by

THEOREM 4.1. FUHRKEN'S FIRST NORMAL FORM THEOREM
A relational structure $\mathfrak{A}$ of cardinal $\mathfrak{m}^+$ is a model of Σ iff there is some subset X of A of cardinal $\leqslant\mathfrak{m}$, and some ternary relation Y on A such that $(\mathfrak{A},X,Y)$ is a model of Σ^*.

PROOF. Suppose first that $\mathfrak{A}$ is a model of Σ. Let X be any subset of A of cardinal $\mathfrak{m}$. Let $\mathscr{A}$ be the collection of all those subsets S of A such that for some subformula $(Qv_n)\,\psi$ of a sentence in Σ, and some sequence $x\in A^\omega$,

$$S=\{a\in A:\mathfrak{A}\vDash_{x(n/a)}\psi\}.$$

Thus $\mathscr{A}$ is the set of those subsets of A definable by subformulas from Σ of the form $(Qv_n)\,\psi$ with parameters from A.

Clearly $\mathrm{card}(\mathscr{A})\leqslant\mathrm{card}(A)$. For each $S\in\mathscr{A}$, let h_S be a one–one map which in case $\mathrm{card}(S)=\mathfrak{m}^+$ maps S onto A and in case $\mathrm{card}(S)\leqslant\mathfrak{m}$ maps S into X, and let

$$H=\{h_S:S\in\mathscr{A}\}.$$

Again $\mathrm{card}(H)\leqslant\mathrm{card}(A)$ and hence we can index H by elements of A. That is, there is a one–one map f of H into A.

We let Y be the ternary relation defined on A by

$$\langle x,y,z\rangle\in Y \text{ iff for some } h\in H,\ f(h)=x \text{ and } h(y)=z.$$

It is clear from the construction of Y that $\mathfrak{B}=(\mathfrak{A},X,Y)$ is a model of those sentences of Σ^* in (b) or (c) above. Since $\mathrm{card}(X)=\mathfrak{m}$, it remains only to show that $\mathfrak{B}$ is also a model of those sentences of Σ^* in (a) above. We will

do this by proving that for all subformulas ϕ of sentences in Σ, and all $x \in A^\omega$,

$$\mathfrak{A} \models_x \phi \quad \text{iff} \quad \mathfrak{B} \models_x \phi^*. \qquad [*]$$

The proof is by induction on the number of logical symbols in ϕ. Trivially [*] holds for atomic formulas and for the formulas $\neg\phi, \phi \wedge \psi, (\exists v_n)\phi$ whenever it is true for ϕ and ψ. So it remains only to show that it is true for $(Qv_n)\phi$ if it is true for ϕ. Assume then that [*] holds for ϕ.

Suppose $\mathfrak{A} \models_x (Qv_n)\phi$. Let

$$S = \{a \in A : \mathfrak{A} \models_{x(n/a)} \phi\},$$

then $\text{card}(S) = \mathfrak{m}^+$. Also $S \in \mathscr{A}$ and so there is some h in H such that h maps S one–one onto A. Let $f(h) = a$. Then for each b in A there is some c in S such that $\langle a,c,b\rangle \in Y$. Also, if $c \in S, \mathfrak{A} \models_{x(n/c)} \phi$ and so, by our hypothesis, $\mathfrak{B} \models_{x(n/c)} \phi^*$. It follows that

$$\mathfrak{B} \models_x (\exists v_k)(\forall v_{k+1})(\exists v_n)[\phi^* \wedge F(v_k, v_n, v_{k+1})],$$

that is

$$\mathfrak{B} \models_x ((Qv_n)\phi)^*.$$

Conversely suppose $\mathfrak{B} \models_x ((Qv_n)\phi)^*$. Then there is some one–one function mapping $S^* = \{a \in A : \mathfrak{B} \models_{x(n/a)} \phi^*\}$ onto A, because this is what $((Qv_n)\phi)^*$ says. By our hypothesis,

$$S^* = S = \{a \in A : \mathfrak{A} \models_{x(n/a)} \phi\}$$

and so $\text{card}(S) = \text{card}(A)$. That is,

$$\mathfrak{A} \models_x (Qv_n)\phi.$$

This completes the proof of [*]. In particular for each sentence σ in $\Sigma, \mathfrak{A} \models \sigma$ and so $(\mathfrak{A}, X, Y) \models \sigma^*$. This concludes the proof that $(\mathfrak{A}, X, Y)$ is a model of Σ^*.

For the converse assume that $\mathfrak{A}$ is a realization of L_Q of cardinal $\mathfrak{m}^+$, X is a subset of A of cardinal $< \mathfrak{m}^+$, and that Y is a ternary relation on A such that $\mathfrak{B} = (\mathfrak{A}, X, Y)$ is a model of Σ^*. We shall prove that $\mathfrak{A}$ is a model of Σ by showing that for any subformula ϕ of some sentence in Σ and any $x \in A^\omega$

$$\mathfrak{B} \models_x \phi^* \quad \text{iff} \quad \mathfrak{A} \models_x \phi. \qquad [+]$$

Again the proof will be by induction on the number of logical symbols in ϕ and once again the only non-trivial step is that for the quantifier Q. Assume then that [+] holds for the formula ϕ; we will show that it holds also for $(Qv_n)\phi$.

Suppose $\mathfrak{B} \vDash_x ((Qv_n)\,\phi)^*$. Then, as before, there is some one–one function mapping $S^* = \{a \in A : \mathfrak{B} \vDash_{x(n/a)} \phi^*\}$ onto A. Again by our induction hypothesis

$$S^* = S = \{a \in A : \mathfrak{A} \vDash_{x(n/a)} \phi\}$$

and so $\mathrm{card}(S) = \mathrm{card}(A)$, whence $\mathfrak{A} \vDash_x (Qv_n)\,\phi$.

Finally suppose $\mathfrak{A} \vDash_x (Qv_n)\,\phi$. Let

$$S = \{a \in A : \mathfrak{A} \vDash_{x(n/a)} \phi\}.$$

Then $\mathrm{card}(S) = \mathrm{card}(A)$, whence as before,

$$\mathfrak{B} \vDash_x (Qv_n)\,\phi^*.$$

$\mathfrak{B}$ is a model of Σ^* and so, in particular,

$$\mathfrak{B} \vDash_{x(k/b)} (\forall v_{k+1})(\exists v_n)[\phi^* \wedge F(v_k, v_n, v_{k+1})] \vee (\forall v_n)(\exists v_{k+1})[\phi^* \to U(v_{k+1}) \wedge F(v_k, v_n, v_{k+1})]$$

for some $b \in A$. Since $\mathfrak{B} \vDash_x (Qv_n)\,\phi^*$ and X, the value of U in $\mathfrak{B}$, is of cardinal $< \mathrm{card}(A)$, it follows that

$$\mathfrak{B} \vDash_{x(k/b)} (\forall v_{k+1})(\exists v_n)[\phi^* \wedge F(v_k, v_n, v_{k+1})],$$

and so

$$\mathfrak{B} \vDash_x (\exists v_k)(\forall v_{k+1})(\exists v_n)[\phi^* \wedge F(v_k, v_n, v_{k+1})],$$

that is

$$\mathfrak{B} \vDash_x ((Qv_n)\,\phi)^*.$$

This completes the proof of $[+]$ and hence of theorem 4.1.

Recall that a set of sentences of L* *admits* the pair of cardinals $\langle \mathfrak{m}, \mathfrak{n} \rangle$, if it has a model of cardinal $\mathfrak{m}$ in which the value of the unary predicate letter U of L* is a set of cardinal $\mathfrak{n}$.

COROLLARY 4.2. *A set Σ of sentences of L_Q has a model of cardinal $\mathfrak{m}^+$ iff Σ^* admits $\langle \mathfrak{m}^+, \mathfrak{n} \rangle$ for some $\mathfrak{n} \leqslant \mathfrak{m}$.*

PROOF. This is an immediate consequence of theorem 4.1.

If Σ is a set of sentences of L_Q we let Σ_α be the set of sentences of L_{Q_α} obtained from L_Q by replacing each occurrence of Q by Q_α. Obviously each set of sentences of L_{Q_α} is of the form Σ_α for some set Σ of sentences of L_Q. Hence with each set Σ_α of sentences of L_{Q_α} we can associate the set Σ^* of sentences of L* as described above. In this way we also obtain the following consequence of theorem 4.1.

COROLLARY 4.3. *A set Σ of sentences of $L_{Q_{\alpha+1}}$ has a model of cardinal $\aleph_{\alpha+1}$ iff Σ^* admits $\langle \aleph_{\alpha+1}, \mathfrak{n} \rangle$ for some $\mathfrak{n} \leqslant \aleph_\alpha$.*

We give some applications of these results in the next section. Meanwhile we consider how the technique we have just described can be extended to limit cardinals. If $\mathfrak{m}$ is a limit cardinal, there is no cardinal $\mathfrak{n} < \mathfrak{m}$ such that any cardinal less than $\mathfrak{m}$ is less than or equal to $\mathfrak{n}$ and so the device used for successor cardinals cannot be applied. In this case we have to use cardinal-like orderings.

We first give a rough sketch of the method, and then describe the formal details. Let $\mathfrak{A}$ be a realization of L of cardinal $\mathfrak{m}$, where $\mathfrak{m}$ is a limit cardinal. We choose a cardinal-like ordering of the domain of $\mathfrak{A}$, say $<$, and for each definable subset X of A we choose a one–one function which either maps X onto A or into an initial segment of A, with respect to the ordering $<$. Since there are only $\mathfrak{m}$ definable subsets of A we can index these one–one functions by elements of A. In this way we define a ternary relation Y on A such that for each $a \in A$ the function g_a given by

$$g_a(b) = c \quad \text{iff} \quad \langle a,b,c \rangle \in Y,$$

is the one–one function mentioned above which is indexed by a. Then instead of saying $(Qv_0)\,\phi$, we say that there is an element of the domain which indexes a function mapping the elements satisfying ϕ onto the whole domain, and instead of saying $\neg(Qv_0)\,\phi$, we say that there is an element which indexes a function mapping the elements satisfying ϕ into an initial segment of the domain. We now give the formal details of this translation.

Let L^+ be the language which is obtained from L by adding a new binary predicate letter P and a new ternary predicate letter F. With each formula ϕ of L_Q we associate the formula ϕ^* of L^+. ϕ^* is defined in the same way as before, when we were dealing with successor cardinals.

Next we associate with each set Σ of sentences of L_Q a set Σ^+ of sentences of L^+. Σ^+ consists of all the following sentences:

(a). All the sentences σ^* for $\sigma \in \Sigma$.

(b). The sentence which says that 'F is an indexed collection of one–one functions' namely

$$(\forall v_0)(\forall v_1)(\exists v_2)(\forall v_3)([F(v_0,v_1,v_3) \rightarrow v_3 = v_2] \wedge (\exists v_2)(\forall v_3)[F(v_0,v_3,v_1) \rightarrow v_3 = v_2]).$$

(c). The sentence which says that P is a linear ordering of the domain, i.e. the sentence σ_0 of example 5.3.6.

(d). For each formula of the form $(Qv_n)\,\psi$ which is a subformula of some sentence in Σ, the closure of

$$(\exists v_k)[(\forall v_{k+1})(\exists v_n)[\psi^* \wedge F(v_k,v_n,v_{k+1})] \vee (\exists v_{k+2})(\forall v_n)(\exists v_{k+1})[\psi^* \to P(v_{k+1},v_{k+2}) \wedge F(v_k,v_n,v_{k+1})]],$$

where v_k is the first variable after v_n and all those in ψ^*.

This is the sentence which says that among the functions indexed by F there is one mapping the elements satisfying ψ^* onto the whole domain or into an initial segment of it with respect to the ordering given by P. We now have

THEOREM 4.4. FUHRKEN'S SECOND NORMAL FORM THEOREM
A relational structure $\mathfrak{A}$ of cardinal $\mathfrak{m}$ is a model of Σ iff there is an $\mathfrak{m}$-like ordering Z of A and a ternary relation Y on A such that $(\mathfrak{A},Z,Y)$ is a model of Σ^+.

PROOF. This theorem is proved in a manner similar to that used to prove Fuhrken's first normal form theorem. The details are left to the reader.

COROLLARY 4.5. *A set Σ of sentences of L_Q has a model of cardinal $\mathfrak{m}$ iff Σ^+ has an $\mathfrak{m}$-like model.*

PROOF. This is an immediate consequence of the previous theorem.

The next section is devoted to some important applications of these results.

§ 5. Löwenheim–Skolem and compactness theorems for L_Q

The results of the previous section and those of ch. 12 enable us to deduce at once the following Löwenheim–Skolem type theorems for L_Q.

THEOREM 5.1. *Let Σ be a countable set of sentences of L_Q. If Σ has a model of cardinal $\aleph_0$ then Σ has a model of each infinite cardinality.*

PROOF. This is an immediate consequence of corollary 4.5 and theorem 12.4.1 which says that a set of sentences with an $\aleph_0$-like model has an $\aleph_\alpha$-like model for each cardinal $\aleph_\alpha$.

THEOREM 5.2. *Let Σ be a countable set of sentences of L_Q. If Σ has a model of cardinal $\aleph_\alpha$ with $\aleph_\alpha$ regular, then Σ has a model of cardinal $\aleph_1$.*

PROOF. This is an immediate consequence of corollary 4.2 and Vaught's two-cardinal theorem.

THEOREM 5.3. (G.C.H.) *Let Σ be a countable set of sentences of L_Q. If for some successor cardinal, say $\aleph_{\beta+1}$, Σ has a model of cardinal $\aleph_{\beta+1}$, then for every regular cardinal $\aleph_\alpha$, Σ has a model of cardinal $\aleph_{\alpha+1}$.*

PROOF. This is an immediate consequence of corollary 4.2 and Chang's two-cardinal theorem.

COROLLARY 5.4. (G.C.H.) *Let Σ be a countable set of sentences of L_Q with a model of some regular cardinality. Then for every regular cardinal $\aleph_\alpha$, Σ has a model of cardinal $\aleph_{\alpha+1}$.*

PROOF. Since $\aleph_1$ is a successor cardinal, this is an immediate consequence of the previous two theorems.

Now that we have these Löwenheim–Skolem theorems for L_Q at our disposal we can use theorem 3.2 to prove the following compactness result.

THEOREM 5.5. (G.C.H.) THE REGULAR COMPACTNESS THEOREM FOR L_Q
Let Σ be a set of sentences of L_Q every finite subset of which has a model of regular cardinality. Then Σ itself has a model.

PROOF. We can assume that Σ is infinite, say $\operatorname{card}(\Sigma)=\aleph_\alpha$. $\aleph_{\alpha+1}$, being a successor cardinal, is certainly regular and therefore, by corollary 5.4, each finite subset of Σ has a model of cardinal $\aleph_{\alpha+2}$. But, by lemma 3.3, since we are assuming the generalized continuum hypothesis, $\aleph_{\alpha+2}$ is $\aleph_\alpha$-normal. The result now follows at once from theorem 3.2.

COROLLARY 5.6. (G.C.H.) *Let Σ be a set of sentences of L_Q of cardinal $\aleph_\alpha$ with a model of regular cardinality. Then Σ has a model of cardinal $\aleph_{\alpha+2}$.*

PROOF. Since in this case every finite subset of Σ has a model of regular cardinality, this is an immediate consequence of the previous result.

Notice that when Σ has a model of some successor cardinal, we could also deduce this corollary from corollary 4.2 and the extension of Chang's two-cardinal theorem to uncountable languages.

A simple example shows that $\aleph_{\alpha+2}$ cannot be replaced by $\aleph_\alpha$ in the previous corollary. Let

$$\Sigma=\{P(c_\xi):\xi<\aleph_\alpha\}\cup\{c_\xi\neq c_\zeta:\xi<\zeta<\aleph_\alpha\}\cup\{\neg(Qv_0)\,P(v_0)\}.$$

It is easily seen that each finite subset of Σ has a model of cardinal $\aleph_0$ and hence of all infinite cardinals whereas Σ itself only has models of cardinals $\geqslant\aleph_{\alpha+1}$.

The question as to whether $\aleph_{\alpha+2}$ can be replaced by $\aleph_{\alpha+1}$ in corollary 5.6 remains open. We conjecture that it can be. In the case $\alpha=0$, this is the content of theorem 5.2.

Our proof of the regular compactness theorem for L_Q involved assuming the generalized continuum hypothesis. It is important to observe that if we restrict our attention to countable sets of sentences then we do not need this hypothesis.

THEOREM 5.7. *Let Σ be a countable set of sentences of L_Q every finite subset of which has a model of regular cardinality, then Σ itself has a model.*

PROOF. Recall first that corollary 4.2 says that a set Δ of sentences of L_Q has a model of cardinal $\aleph_1$ iff Δ^* admits $\langle \aleph_1, \aleph_0 \rangle$. Also notice that it follows from lemmas 12.6.2 and 12.6.3 that Δ^* admits $\langle \aleph_1, \aleph_0 \rangle$ iff $(\Delta^*)_W$ is consistent.

Now suppose Σ_0 is a finite subset of Σ. Since Σ_0 has a model of regular cardinality, by theorem 5.2, Σ_0 has a model of cardinal $\aleph_1$. Therefore $(\Sigma_0{}^*)_W$ is consistent. Now for each finite subset Δ_0 of $(\Sigma^*)_W$ there is some finite subset Σ_0 of Σ such that

$$\Delta_0 \subseteq (\Sigma_0{}^*)_W .$$

Therefore each finite subset of $(\Sigma^*)_W$ is consistent. Hence $(\Sigma^*)_W$ is itself consistent and we can therefore conclude that Σ has a model of cardinal $\aleph_1$. This completes the proof.

In exactly the same way we obtain

THEOREM 5.8. FUHRKEN'S COMPACTNESS THEOREM
The language L_{Q_1} is $\aleph_0$-compact.

PROOF. The proof of this result follows that of the previous theorem, except that the downward Löwenheim–Skolem theorem for L_{Q_α} is used instead of theorem 5.2, and corollary 4.3 instead of corollary 4.2.

In order to be able to settle whether or not the compactness theorem in its full generality holds for L_Q we would need to know what happens in the case of sets of sentences with models of singular cardinality. Some recent work of Keisler provides the answers to these questions. Unfortunately it is not possible to explain Keisler's methods here. Essentially they depend on a translation like that used to obtain theorem 4.4. But Keisler's translation is much more complicated, having affinities with the work of SILVER [1966]. Also in order to show that his translation works, Keisler needs to use a complicated set theoretic argument which is itself based on the polarized partition relations for cardinals of ERDÖS, HAJNAL and RADO [1965]. This will explain why we do not describe Keisler's method. We content ourselves with just stating his results.

Recall that a cardinal κ is said to be a *strong limit cardinal* if for all cardinals $\alpha < \kappa$, $2^{\alpha} < \kappa$.

THEOREM 5.9. *Let κ be a strong limit cardinal and let λ be a singular cardinal. If Σ is a countable set of sentences of L_Q with a model of cardinal κ, then Σ also has a model of cardinal λ.*

THEOREM 5.10. *Let Σ be a set of sentences of L_Q each finite subset of which has a model of some singular strong limit cardinal. Then Σ itself has a model.*

Theorem 5.9 can be extended to uncountable sets of sentences providing that we insist that κ and λ are both greater than $\text{card}(\Sigma)$. Proofs of these theorems are given in KEISLER [1968i].

In order to put theorems 5.5 and 5.10 together we shall have to make the assumption that all *singular cardinals are strong limit cardinals*. We call this assumption the limit cardinal hypothesis (abbreviated as L.C.H.). Since all singular cardinals are limit cardinals the L.C.H. is implied by the generalized continuum hypothesis, and therefore, by the results of GÖDEL [1940], the L.C.H. is consistent with set theory.

Another way of looking at the L.C.H. is the following. Let f be the ordinal valued function of ordinals such that

$$2^{\aleph_\xi} = \aleph_{\xi + f(\xi)}.$$

Cantor's theorem on the cardinality of power sets tells us that f never takes the value 0. The generalized continuum hypothesis is equivalent to saying that f always takes the value 1. The limit cardinal hypothesis is equivalent to saying that f is always finite (so that the operation of forming the power set never takes you as far as the next limit cardinal). The results of EASTON [1964] show that it is consistent with set theory to choose f in a large number of ways. In particular it follows from his results that the L.C.H. is strictly weaker than the generalized continuum hypothesis, but that it is not provable in set theory with the axiom of choice.

Theorems below that depend on the L.C.H. will be indicated by putting '(L.C.H.)' after the numbers of the relevant theorems.

THEOREM 5.11. (L.C.H.) THE SINGULAR COMPACTNESS THEOREM FOR L_Q *Let Σ be a set of sentences of L_Q each finite subset of which has a model of singular cardinality. Then Σ itself has a model.*

PROOF. Immediate from theorem 5.10.

Putting these results together we now obtain

THEOREM 5.12. (G.C.H.) THE COMPACTNESS THEOREM FOR L_Q
A set of sentences Σ of L_Q has a model iff each finite subset of Σ has a model.

PROOF. One way round is trivial. For the other suppose that each finite subset of Σ has a model. Then either

(1). each finite subset of Σ has a model of regular cardinality,

or

(2). each finite subset of Σ has a model of singular cardinality.

For if there is a finite subset Σ_0 of Σ with no model of regular cardinality and a finite subset Σ_1 of Σ with no model of singular cardinality, then $\Sigma_0 \cup \Sigma_1$ is a finite subset of Σ with no model, contradicting our assumption.

If (1) holds the result follows from the regular compactness theorem for L_Q and if (2) holds it is a consequence of the singular compactness theorem for L_Q. This completes the proof.

We conclude this section by remarking that using theorem 5.7 it can be seen that the $\aleph_0$-compactness of L_Q can be proved assuming only the L.C.H. and not the G.C.H.

§ 6. Completeness theorem for L_Q

As is well known the question as to whether there exists a recursive set of axioms sufficient to generate all the universally valid sentences of some language L*, is equivalent to the question whether the universally valid sentences of L* form a recursively enumerable set (all this, of course, under the assumption that we have some effective Gödel numbering for the formulas of L*). We begin by discussing this second question, and only towards the end of this section do we turn our attention to the problem of actually finding an explicit axiomatization for L_Q.

We let *Val* be the set of universally valid sentences of L_Q. For each ordinal α, Val_α is the set of sentences of L_Q true in all realizations of L_Q of cardinal $\aleph_\alpha$. Val_R will denote the set of these sentences true in all realizations of L_Q of regular cardinality and Val_S will denote the set of those sentences of L_Q true in all realizations of singular cardinality. Thus

$$Val = Val_R \cap Val_S = \bigcap_\alpha Val_\alpha,$$

where the last intersection is taken over all ordinals.

THEOREM 6.1. *Val_1 is recursively enumerable.*

PROOF. A sentence σ of L_Q is true in all realizations of cardinal $\aleph_1$ iff $\neg\sigma$ has no models of cardinal $\aleph_1$. By corollary 4.2, this holds iff $\{\neg\sigma\}^*$ does not admit $\langle\aleph_1,\aleph_0\rangle$. Notice now that $\{\neg\sigma\}^*$ is a finite set of sentences and that given σ we can effectively find σ', the conjunction of all the sentences in $\{\neg\sigma\}^*$.

Now, by lemmas 12.6.2 and 12.6.3, $\{\sigma'\}$ fails to admit $\langle\aleph_1,\aleph_0\rangle$ iff $\{\sigma'\}_W$ is not consistent. Next observe that

$$\{\sigma'\}_W = \{\sigma'\} \cup \Delta,$$

where Δ, which is independent of σ', consists of the sentences

$$(\forall v_0)[U(v_0) \to W(v_0)], \quad (\exists v_0)\neg W(v_0),$$

and all the sentences ϕ_W, where ϕ is a formula of L.

Thus $\{\sigma'\}_W$ is not consistent iff $\Delta \vdash \neg\sigma'$. Hence we have shown that

$$\sigma \in Val_1 \quad \text{iff} \quad \Delta \vdash \neg\sigma'. \tag{1}$$

Δ is a recursive set of sentences and σ' can be found effectively given σ. Also they are all sentences of L, which is recursively axiomatizable. It therefore follows at once from (1) that Val_1 is recursively enumerable.

In a similar way we can prove an analogous result for L_{Q_1}. This result, which is due to Vaught, is the original version of the theorem we have just proved.

THEOREM 6.2. VAUGHT'S COMPLETENESS THEOREM FOR L_{Q_1}
The set of universally valid sentences of L_{Q_1} is recursively enumerable.

PROOF. The proof of this theorem is similar to that of the previous one, using the downward Löwenheim–Skolem theorem for L_{Q_1}.

KEISLER [1970] has shown that the following axiom schemas, when added to those for first order predicate calculus, form a complete axiomatization for L_{Q_1}, in the sense that any set of sentences of L_{Q_1} consistent with them has a model.

K1. $(\forall y)(\forall z)\neg(Q_1x)[x=y \vee x=z]$,
K2. $(\forall x)(\phi\to\psi)\to((Q_1x)\,\phi\to(Q_1x)\,\psi)$,
K3. $(Q_1x)\,\phi(x)\leftrightarrow(Q_1y)\,\phi(y)$,
K4. $(Q_1y)(\exists x)\,\phi\to[(Q_1x)(\exists y)\,\phi \vee (\exists x)(Q_1y)\,\phi]$,

where in K3, y does not occur in $\phi(x)$. Keisler's work yields alternative proofs of theorems 5.2, 6.1 and 6.2.

Using the Löwenheim–Skolem theorems for L_Q obtained in the previous section we obtain

THEOREM 6.3. (G.C.H.) *If $\aleph_\alpha$ is a regular cardinal then*

$$Val_{\alpha+1} = Val_1,$$

and is therefore recursively enumerable.

PROOF. $\aleph_{\alpha+1}$ is certainly regular and therefore, by theorem 5.2,

$$Val_1 \subseteq Val_{\alpha+1},$$

while, by theorem 5.3,

$$Val_{\alpha+1} \subseteq Val_1,$$

and this proves the theorem.

From theorem 6.1 we can deduce that Val_R *is recursively enumerable.*

$$Val_R = \bigcap \{Val_\alpha : \aleph_\alpha \text{ is regular}\}$$

and so certainly $Val_R \subseteq Val_1$. But by theorem 5.2, if σ is true in all models of cardinal $\aleph_1$ then σ is true in all models of regular cardinality. Thus $Val_1 \subseteq Val_R$, whence

$$Val_R = Val_1,$$

which we have just shown to be recursively enumerable.

If in K1–K4, 'Q_1' is replaced by 'Q', we obtain a set of axioms whose consequences are precisely the sentences in Val_R.

From theorem 5.9 we can deduce, in a similar way, that assuming the L.C.H.,

$$Val_S = Val_\omega$$

Another result due to Keisler now enables us to derive a completeness theorem for L_Q. Once again we only quote his result.

THEOREM 6.4. *If $\aleph_\alpha$ is a singular strong limit cardinal, then Val_α is recursively enumerable.*

The proof of this theorem can also be found in KEISLER [1968i].

Now

$$Val = Val_R \cap Val_S = Val_1 \cap Val_\omega,$$

and by the previous results of this section both Val_1 and Val_ω are recursively enumerable. The intersection of two recursively enumerable sets is itself a recursively enumerable set and so we have shown

THEOREM 6.5. (L.C.H.) THE COMPLETENESS THEOREM FOR L_Q
The set of universally valid sentences of L_Q is recursively enumerable.

It is natural to enquire whether this completeness theorem could be proved without assuming the L.C.H. This question remains open. We remark here that precisely which sentences of L_Q are universally valid sentences is not independent of assumptions about cardinal exponentiation. That is we can find a sentence of L_Q whose universal validity is independent of set theory with the axiom of choice. Indeed as we now show we can find a sentence of L_Q whose universal validity is equivalent to the limit cardinal hypothesis.

Let σ be the sentence

$$\neg(Qv_0)(\exists v_1)\ R(v_0,v_1) \wedge (\forall v_0)\neg(Qv_1)\ R(v_0,v_1) \wedge (Qv_1)(\exists v_0)\ R(v_0,v_1)$$

and suppose $\mathfrak{A}=\langle A,R_{\mathfrak{A}}\rangle$ is a model of σ with $\mathrm{card}(A)=\aleph_\alpha$. To each element $a\in A$ there corresponds the set

$$X_a=\{b\in A:\mathfrak{A}\models R[a,b]\}$$

of elements of A. The first conjunct of σ says that fewer than $\aleph_\alpha$ of these sets are non-empty, the second conjunct that each set X_a contains fewer than $\aleph_\alpha$ elements and the third conjunct that the union of the sets X_a is of cardinal $\aleph_\alpha$. Clearly this implies that $\aleph_\alpha$ is singular. Conversely it is easy to see that if $\aleph_\alpha$ is singular, then σ has a model of cardinal $\aleph_\alpha$. That is

LEMMA 6.6. *σ has a model of cardinal $\aleph_\alpha$ iff $\aleph_\alpha$ is singular.*

On the other hand we can find a sentence τ of L_Q which only has models of non-strong limit cardinals. τ is the sentence

$$(\forall v_0)(\forall v_1)[(\forall v_2)[P(v_0,v_2)\leftrightarrow P(v_1,v_2)]\rightarrow v_0=v_1]\wedge\neg(Qv_2)(\exists v_0)\ P(v_0,v_2).$$

Again if $\mathfrak{A}=\langle A,P_{\mathfrak{A}}\rangle$ is a model of τ with $\mathrm{card}(A)=\aleph_\alpha$, to each element $a\in A$ there corresponds the set

$$X_a=\{b\in A:\mathfrak{A}\models P[a,b]\}$$

of elements of A. The first conjunct of τ says that if $a\neq b$ then $X_a\neq X_b$ and the second that the union of the sets X_a, say X, is of cardinal less than $\aleph_\alpha$, say $\aleph_\beta$ with $\beta<\alpha$. There are at most $2^{\aleph_\beta}$ distinct sets made up of elements of X and hence $\aleph_\alpha\leqslant 2^{\aleph_\beta}$, or else we would not be able to associate distinct subsets of X with distinct elements of A. Thus $\aleph_\alpha$ is not a strong limit cardinal. Conversely it is easy to see that if $\aleph_\alpha$ is not a strong limit cardinal, then τ has a model of cardinal $\aleph_\alpha$. Thus we have proved

LEMMA 6.7. *τ has a model of cardinal $\aleph_\alpha$ iff $\aleph_\alpha$ is not a strong limit cardinal.*

The last two lemmas together show that *the sentence*

$$\neg(\sigma \wedge \tau)$$

is universally valid iff the limit cardinal hypothesis holds. Thus although it may be possible to show that L_Q is recursively axiomatizable without assuming the L.C.H., it is not possible to provide an explicit axiomatization without making some assumption about cardinal exponentiation which is independent of set theory with the axiom of choice. From this it would seem that the problem of finding a 'nice' explicit recursive axiomatization for L_Q is very difficult and not much progress has been made towards finding one. YASUHARA [1966] has provided a Gentzen type axiomatization for the fragment L_Q^- of L_Q *without* equality. Yasuhara's system is equivalent to that obtained from our system of axioms for predicate calculus with equality by dropping the axioms involving equality and adding the following axiom schemas:

Q1. $(\forall x)\,\phi(x) \rightarrow (Qx)\,\phi(x)$,
Q2. $(Qx)\,\phi(x) \rightarrow (\exists x)\,\phi(x)$,
Q3. $(Qx)[\phi(x) \vee \psi(x)] \rightarrow (Qx)\,\phi(x) \vee (Qx)\,\psi(x)$,
Q4. $(Qx)\,\phi(x) \wedge (\forall x)[\phi(x) \rightarrow \psi(x)] \rightarrow (Qy)\,\psi(y)$,

where x,y are any variables and ϕ,ψ any formulas of L_Q^-. It should be noted that Q3 is only valid for infinite realizations. The change of variable in Q4 is necessary for us to be able to prove the equivalence

$$(Qx)\,\phi(x) \leftrightarrow (Qy)\,\phi(y)$$

which would not otherwise be provable.

Yasuhara showed that a countable set of sentences of L_Q^- which is consistent with respect to this axiomatization has a model of singular cardinality. This result can easily be extended to uncountable sets of sentences to show that a set of sentences of L_Q^- of cardinal κ which is consistent has a model of all singular cardinals whose cofinality is not less than κ.

Now let L_Q^1 be the fragment of L_Q containing equality but otherwise only unary predicate letters. It is shown in SLOMSON [1968] that if we add to our axiom system for predicate calculus with equality the four schemas above and the single schema

$$(\forall y)\neg(Qx)(x=y),$$

where x,y are distinct variables, then any set of sentences of L_Q^1 consistent with respect to these axioms has a model.

Thus explicit axiomatizations have been provided for both L_Q^- and L_Q^1.

These fragments of L_Q have in common the fact that it is not possible to express the notion of a function in either of them. But if we have a quantifier like Q in the language we want to be able to say things about one–one functions and so the fragments L_Q^- and L_Q^1 are not very interesting. Clearly any complete axiomatization for L_Q itself will include an axiom which says that 'large' sets are preserved under one–one functions. The remarks above show that if we assume the L.C.H., another axiom would be required to ensure that $\neg(\sigma \wedge \tau)$ is provable. It is tempting to conjecture that by adding an axiom about one–one functions and the axiom $\neg(\sigma \wedge \tau)$ to the axiomatization for L_Q^1 given above we would obtain a complete axiomatization for L_Q. But a proof of this, even if it were true, would appear to be difficult to find.

We have now seen that L_Q is an extension of first order predicate calculus with equality which, assuming the L.C.H., is recursively axiomatizable and, assuming the G.C.H., satisfies the compactness theorem. Also several Löwenheim–Skolem type theorems hold for L_Q. We conclude this chapter with two further remarks which indicate that L_Q is a weak extension of L.

First suppose that $\mathfrak{A}=\langle R, < \rangle$ is the set of rationals with their usual ordering and that $\mathfrak{B}=\langle S, < \rangle$ is the set of real numbers with their usual ordering. We have already seen (ch. 4), that with respect to the language L, $\mathfrak{A}$ is elementarily equivalent to $\mathfrak{B}$, so that L is not powerful enough to distinguish between the order type of the rationals and that of the reals. The same is true for the language L_Q. We leave the proof of this to the reader; it can be proved using a method similar to that we used to prove the analogous result for L.

Finally, suppose that *FA* is any recursive set of axioms for formal arithmetic in the language L. It follows from Gödel's incompleteness theorem (GÖDEL [1931]) that *FA* is not complete. Let

$$FA_Q = FA \cup \{(\forall x)\neg(Qy)\ y < x\}.$$

It easily seen that FA_Q is an $\aleph_0$-categorical set of axioms for arithmetic. An argument similar to that used to prove theorem 2.8 shows that FA_Q is also not complete. But it might be the case that there are sentences of L provable from FA_Q which are not provable from *FA*. The next result shows that this is not so, so that FA_Q is a conservative extension of *FA*.

THEOREM 6.8. *Let σ be a sentence of L which is provable from FA_Q, then σ is also provable from FA.*

PROOF. Suppose not $FA \vdash \sigma$. Then $FA \cup \{\neg\sigma\}$ is consistent and therefore

has a countable model say $\mathfrak{A}$. We define the sequence $\langle \mathfrak{A}_\xi : \xi < \aleph_1 \rangle$ of countable models of $FA \cup \{\neg\sigma\}$ as follows:

$$\mathfrak{A}_0 = \mathfrak{A}.$$

If $\lambda < \aleph_1$ is a limit ordinal

$$\mathfrak{A}_\lambda = \bigcup_{\xi<\lambda} \mathfrak{A}_\xi$$

and if $\xi = \zeta + 1$ is a successor ordinal $< \aleph_1$ then $\mathfrak{A}_\xi$ is the countable proper elementary end extension of $\mathfrak{A}_\zeta$ whose existence is guaranteed by MacDowell and Specker's theorem. Then

$$\mathfrak{A}^* = \bigcup_{\xi<\aleph_1} \mathfrak{A}_\xi$$

is an $\aleph_1$-like model of $FA \cup \{\neg\sigma\}$. Therefore it is also a model of $FA_Q \cup \{\neg\sigma\}$. Therefore $FA_Q \cup \{\neg\sigma\}$ is consistent and so not $FA_Q \vdash \sigma$. This completes the proof.

§ 7. Historical and bibliographical remarks

The idea of considering generalized quantifiers is due to MOSTOWSKI [1957]. Theorems 2.6 and 2.8 as well as many of the remarks in § 1 are taken from this paper. Theorems 2.2 and 2.3 are due to FUHRKEN [1964]. The results of § 3 come from SLOMSON [1968] as also does theorem 5.5 and the remarks at the end of § 5. Fuhrken's normal form theorems are proved in FUHRKEN [1964], where theorems 5.1 and 5.2 can be found. Theorem 5.3 comes from FUHRKEN [1965]. Throughout Fuhrken states his results in terms of the quantifiers Q_α. Theorem 6.2 is due to VAUGHT [1964].

CHAPTER 14

INFINITARY LANGUAGES

All the languages considered up to this point have one salient feature in common, namely, in each case the class of well formed formulas is obtained from the class of atomic formulas by closing up under *finite* applications of the connectives and quantifiers, so that the resulting expressions are all of finite length. In this chapter we consider languages whose syntactic rules admit the formation of *infinitely* long expressions.

§1. The language $L_{\alpha\beta}$

We define below, for each pair of infinite cardinals α,β with $\alpha \geqslant \beta$, an infinitary language $L_{\alpha\beta}$ whose rules allow conjunctions and disjunctions of length $<\alpha$ and quantifications over strings of variables of length $<\beta$.

The expressive power of an infinitary language will, in general, be incomparably greater than that of any first order language. For instance, we have already remarked that infinite structures cannot be characterized by categorical axioms. In particular, we noticed in ch. 12 that we cannot even find an $\aleph_0$-categorical set of axioms which describe the standard model of arithmetic. In contrast, however, this model *can* be characterized up to isomorphism by the single sentence

$$\sigma = \sigma_1 \wedge \sigma_2 \wedge \sigma_3 \wedge \sigma_4$$

of $L_{\omega_1\omega}$, where

$$\sigma_1 = (\forall x) \neg S(x,0),$$
$$\sigma_2 = (\forall xyz)[S(x,z) \wedge S(y,z) \rightarrow x = y],$$
$$\sigma_3 = (\forall x)(\exists y)(\forall z)[S(x,z) \leftrightarrow y = z],$$

and

$$\sigma_4 = (\forall x)[\psi_0 \vee \psi_1 \vee \cdots \vee \psi_n \vee \cdots],$$

where, for each $n<\omega$, ψ_n is the formula

$$(\exists y_0 \ldots y_n)[y_0 = 0 \wedge S(y_0,y_1) \wedge \cdots \wedge S(y_{n-1},y_n) \wedge y_n = x]$$

As a further example we observe that the class of relations which are well-

orderings can be characterized by a single $L_{\omega_1\omega_1}$ sentence, namely

$$(\forall xyz)[R(x,y) \wedge R(y,z) \rightarrow R(x,z)] \wedge (\forall xy)[R(x,y) \vee x = y \vee R(y,x)]$$
$$\wedge (\forall x_1 \ldots x_n \ldots) \neg [R(x_2,x_1) \wedge R(x_3,x_2) \wedge \cdots \wedge R(x_{n+1},x_n) \wedge \cdots],$$

whereas we have already shown (lemma 7.1.13) that there is no set of sentences of any first order language which characterizes well-orderings; indeed we showed that the collection of well-ordered sets is not even an $EC_{\Delta\Sigma}$.

This example indicates that $L_{\omega_1\omega_1}$ behaves rather like a second order language, and since in it we can make statements like 'for all countable subsets X...' and 'there exists a countable subset Y such that ...' by quantifying over a countable sequence of variables, $L_{\omega_1\omega_1}$ may be regarded as a second order language with respect to countable structures. In view of this, it is not too surprising that some of the standard metatheorems of first order logic fail to hold for $L_{\omega_1\omega_1}$ (and, more generally, for $L_{\alpha\alpha}$ with $\alpha \geqslant \omega_1$). Our main concern in this chapter will be to show that one of these metatheorems, the compactness theorem, fails to generalize to $L_{\omega_1\omega_1}$ and to most of the languages $L_{\alpha\alpha}$, with α uncountable, even when α is inaccessible. Thus, in a sense, the infinitary languages $L_{\alpha\alpha}$ are too powerful to admit a reasonable model theory, or, at any rate, a model theory obtained as a simple minded extension of first order model theoretic ideas.

We now proceed to an explicit formulation of the language $L_{\alpha\beta}(\mu)$, where α and β are infinite cardinals with $\alpha \geqslant \beta$ and, for some ordinal ρ, $\mu \in \beta^\rho$. (More often than not we shall not bother to mention the type μ explicitly and we just write $L_{\alpha\beta}$.) The language $L_{\alpha\beta}(\mu)$ has the following basic symbols:

individual variables	v_ξ for $\xi < \alpha$
equality symbol	$=$
predicate letters	P_ξ for $\xi < \rho$
finitary connectives	$\neg$ and $\wedge$
infinitary connective	$\bigwedge$
existential quantifier	$\exists$

Atomic formulas are expressions of the form $v_\xi = v_\zeta$ or $P_\xi(\boldsymbol{v})$, where $\boldsymbol{v}$ is a sequence of variables of length $\mu(\xi)$. The class of (well-formed) *formulas* is defined recursively as follows:

(1). Each atomic formula is a formula.

(2). If ϕ and ψ are formulas, so also are $\neg\phi$ and $\phi \wedge \psi$.

(3). If $\xi < \alpha$ and for $\zeta < \xi$, ϕ_ζ is a formula then

$$\bigwedge_{\zeta<\xi} \phi_\zeta$$

is also a formula.

(4). If ϕ is a formula and v is a sequence of variables of length $<\beta$, then $(\exists v)\phi$ is also a formula.

(5). An expression is a formula only if it follows from a finite number of applications of clauses (1)–(4) that it is a formula.

The symbols $\bigvee$ for infinitary disjunction and $\forall$ for universal quantification are defined in the obvious way as abbreviations, that is

$$\bigvee_{\zeta<\xi}\phi_\zeta \quad \text{abbreviates} \quad \neg\bigwedge_{\zeta<\xi}\neg\phi_\zeta,$$

and

$$(\forall v)\,\phi \quad \text{abbreviates} \quad \neg(\exists v)\neg\phi.$$

We also use the symbols $\vee, \rightarrow$ and $\leftrightarrow$ as abbreviations in the usual way.

Notice that it follows from these definitions that $L_{\omega\omega}(\mu)$ is simply the first order language appropriate for structures of type μ.

As *realizations* of the language $L_{\alpha\beta}(\mu)$ we take structures of the form

$$\langle A,\{R_\xi:\xi<\rho\}\rangle,$$

consisting of a non-empty set A and a collection $\{R_\xi:\xi<\rho\}$ of relations, where for $\xi<\rho$,

$$R_\xi\subseteq A^{\mu(\xi)},$$

that is R_ξ is a $\mu(\xi)$-ary relation on A. We call such structures *μ-structures* (or, occasionally, *$L_{\alpha\beta}$-structures*).

The notion of a valuation of the free variables *satisfying* a formula of $L_{\alpha\beta}$ in a given μ-structure can now be formulated in the customary manner. There is one slight technical difficulty. In formulas of the sort $(\exists v)\,\phi$ we have allowed the case where the sequence v contains more than one occurrence of the same variable. Thus when we relate the truth of $(\exists v)\,\phi$ to the satisfiability of ϕ we have to ensure that we interpret the same variables always as the same elements. This prompts the following definitions.

Suppose A is a set, ξ is an ordinal less than β, b is a ξ-termed sequence of elements of A, say $b=\langle b_\zeta:\zeta<\xi\rangle$ and τ is a ξ-termed sequence of ordinals less than α, that is $\tau\in\alpha^\xi$. We say that the sequence b is *τ-acceptable* if whenever, for $\zeta,\zeta'<\xi$, $\tau(\zeta)=\tau(\zeta')$, also $b_\zeta=b_{\zeta'}$. If a is a valuation of the variables of $L_{\alpha\beta}$, say $a=\langle a_\eta:\eta<\alpha\rangle\in A^\alpha$ and b is τ-acceptable we define the sequence $a[\tau,b]\in A^\alpha$ as follows:

$a[\tau,b]_\eta=a_\eta$, if η is not in the range of τ,

$a[\tau,b]_\eta=b_\zeta$, if η is in the range of τ, say $\tau(\zeta)=\eta$.

The condition that b is τ-acceptable is just the condition we need to ensure that this definition is unambiguous. $a[\tau,b]$ is therefore the sequence which

agrees with a except at places corresponding to the ordinals in the range of τ, when the values are given by the sequence b.

Now suppose $\mathfrak{A}=\langle A,\{R_\xi:\xi<\rho\}\rangle$ is a μ-structure and

$$a=\langle a_\xi:\xi<\alpha\rangle\in A^\alpha$$

is a valuation of the variables of $L_{\alpha\beta}$. We define the relation *a satisfies the formula ϕ in $\mathfrak{A}$*, which we write as $\mathfrak{A}\models_a\phi$, recursively as follows:

(1). $\mathfrak{A}\models_a v_\xi=v_\zeta$ iff a_ξ is the same element as a_ζ,

(2). $\mathfrak{A}\models_a P_\xi(v)$ iff $\langle a_{\xi_\eta}:\eta<\mu(\xi)\rangle\in R_\xi$, where $v=\langle v_{\xi_\eta}:\eta<\mu(\xi)\rangle$,

(3). $\mathfrak{A}\models_a\neg\phi$ iff not $\mathfrak{A}\models_a\phi$,

(4). $\mathfrak{A}\models_a\phi\wedge\psi$ iff $\mathfrak{A}\models_a\phi$ and $\mathfrak{A}\models_a\psi$,

(5). $\mathfrak{A}\models_a\bigwedge_{\eta<\xi}\phi_\eta$ iff for all $\eta<\xi$, $\mathfrak{A}\models_a\phi_\eta$.

(6). If $v=\langle v_{\tau(\eta)}:\eta<\xi\rangle$,

$\mathfrak{A}\models_a(\exists v)\,\phi$ iff for some τ-acceptable sequence $b\in A^\xi$, $\mathfrak{A}\models_{a[\tau,b]}\phi$.

It can be seen that the first four of these clauses are just the same as those of our satisfaction definition for L and the other two are just what would be expected from our intuitive understanding of infinite conjunctions and quantifications.

We can now extend to $L_{\alpha\beta}$ the usual model theoretic definitions of *valid*, *model* etc. in the obvious way. For each set Σ of sentences of $L_{\alpha\beta}$ we let $Mod(\Sigma)$ be the collection of all μ-structures which are models of Σ, and if K is a class of μ-structures, we let $Th(K)$ be the set of sentences of $L_{\alpha\beta}$ which are true in all structures in K.

§ 2. The compactness property; incompactness results for accessible cardinals

We have already seen in ch. 5 that first order predicate logic has the *compactness property*, that is, a set of first order sentences has a model if and only if each finite subset has a model. This notion can be extended to infinitary languages $L_{\alpha\beta}$ in the following way. Let us say that a set of sentences Σ of $L_{\alpha\beta}$ is *locally consistent* if every subset of Σ of cardinality $<\alpha$ has a model. Then the language $L_{\alpha\beta}$ is said to have the *compactness property* if every locally consistent set of sentences of $L_{\alpha\beta}$ has a model. Our object in this and the next section is to show that the vast majority of languages $L_{\alpha\alpha}$ with $\alpha>\omega$ fail to have this property. Indeed, for all strongly accessible cardinals α, and most inaccessible cardinals $\alpha>\omega$ we shall exhibit a locally consistent set of sentences of $L_{\alpha\alpha}$ of cardinal α with no model.

Let α and β be infinite cardinals. α is said to be *β-incompact* if there is a locally consistent set of sentences of $L_{\alpha\alpha}$ of cardinal β which has no

model. We shall call α *incompact* if it is β-incompact for some $\beta \geqslant \alpha$ (clearly α is not β-incompact if $\beta < \alpha$) and *strongly incompact* if it is α-incompact. A cardinal which is not incompact will be said to be *compact* and one that is not strongly incompact will be said to be *weakly compact*. By the compactness theorem for first order languages, $\aleph_0$ is compact.

THEOREM 2.1. *If $\alpha \leqslant \beta^\gamma$ with $\beta, \gamma < \alpha$ then α is β^γ-incompact.*

PROOF. Consider the language $L_{\alpha\alpha}$ with the set

$$\{P_{\xi\eta} : \xi < \gamma \text{ and } \eta < \beta\}$$

of proposition letters (i.e. zero-place predicate letters). Let Σ be the set of sentences of $L_{\alpha\alpha}$ consisting of the sentence

$$\bigwedge_{\xi<\gamma} \bigvee_{\eta<\beta} P_{\xi\eta}, \tag{1}$$

and all the sentences

$$\neg \bigwedge_{\xi<\gamma} P_{\xi f(\xi)}, \tag{2}$$

with $f \in \beta^\gamma$.

We claim that Σ has no model. For suppose $\mathfrak{A}$ is a model of Σ. Since $\mathfrak{A}$ is a model of (1) for each $\xi < \gamma$ there is some $\eta < \beta$ such that $P_{\xi\eta}$ holds in $\mathfrak{A}$. If $\xi < \gamma$, let $f(\xi)$ be the least such η. Thus $f \in \beta^\gamma$ is a function such that $\mathfrak{A}$ satisfies

$$\bigwedge_{\xi<\gamma} P_{\xi f(\xi)}.$$

This contradicts the fact that $\mathfrak{A}$ is a model of all the sentences (2) and this contradiction shows that Σ has no model.

On the other hand it is not difficult to see that Σ is locally consistent. For suppose that Σ' is a subset of Σ of cardinal $< \alpha$ (or indeed any subset of Σ which does not contain all the sentences (2)). Then there is some $f \in \beta^\gamma$ such that the corresponding sentence (2) is not in Σ'. If we now define $\mathfrak{A}$ by letting $P_{\xi\eta}$ be true in $\mathfrak{A}$ whenever $\eta = f(\xi)$ and false otherwise, it is clear that $\mathfrak{A}$ is a model of Σ'.

Now Σ is of cardinal β^γ; it follows therefore that α is β^γ-incompact.

Before proceeding to the next result we fix some notation. We let P_0 be a fixed binary predicate letter and for convenience we write $x < y$ for $P_0(x, y)$. Let σ_0 be the sentence

$$(\forall uvw)[u < v \wedge v < w \rightarrow u < w] \wedge (\forall uv)[u < v \vee u = v \vee v < u],$$

and let σ_{WO} be the sentence

$$\sigma_0 \wedge \neg(\exists \langle v_n : n < \omega \rangle)[\bigwedge_{n<\omega} (v_{n+1} < v_n)].$$

σ_{WO} is a sentence of each of the languages $L_{\alpha\alpha}$ with $\alpha > \omega$. Clearly if $\mathfrak{A} = \langle A, R_0, \ldots \rangle$ is a model of σ_{WO} then R_0 is a well-ordering of the set A. Thus such a model $\mathfrak{A}$ is isomorphic to a structure $\langle \sigma, <, \ldots \rangle$ where σ is an ordinal and $<$ is the 'less than' relation restricted to σ. Hence when we are considering models of σ_{WO} we can restrict our attention to structures of this form.

It will also be convenient to have for each ordinal η a formula $Or_\eta(v_0)$ characterizing the ordinal η. We take for $Or_\eta(v_0)$ the formula

$$(\exists \langle u_\xi : \xi < \eta \rangle)\Big[\bigwedge_{\xi<\eta}\bigwedge_{\zeta<\xi}(u_\zeta < u_\xi) \wedge (\forall w)\Big[w < v_0 \rightarrow \bigvee_{\xi<\eta}(w = u_\xi)\Big]\Big].$$

It is clear that $Or_\eta(v_0)$ is a formula of $L_{\alpha\alpha}$ whenever $\eta < \alpha$ and that in a structure $\langle \sigma, <, \ldots \rangle$ an ordinal τ satisfies $Or_\eta(v_0)$ if and only if $\tau = \eta$.

We are now ready for the next result.

THEOREM 2.2. *Each singular cardinal is strongly incompact.*

PROOF. Suppose α is a singular cardinal. Then for some cardinal $\beta < \alpha$ there is a collection $\{\gamma_\xi : \xi < \beta\}$ of cardinals each less than α such that

$$\alpha = \sum_{\xi<\beta} \gamma_\xi.$$

Let P_1 be a unary predicate letter and let Σ be the set of sentences of $L_{\alpha\alpha}$ consisting of the sentence σ_{WO}, the sentence

$$\bigvee_{\xi<\beta}(\exists \langle v_\eta : \eta < \gamma_\xi \rangle)(\forall v)\Big[P_1(v) \rightarrow \bigvee_{\eta<\gamma_\xi}(v = v_\eta)\Big], \tag{1}$$

and, for each $\eta < \alpha$, the sentence

$$(\exists v)[Or_\eta(v) \wedge P_1(v)]. \tag{2}$$

We show first that Σ has no model. Suppose, on the contrary, that $\mathfrak{A}$ is a model of Σ. Then $\mathfrak{A}$ is isomorphic to a structure $\langle \sigma, <, R_1, \ldots \rangle$ where R_1 is a subset of σ. We can suppose that $\mathfrak{A}$ is actually of this form. Since $\mathfrak{A}$ is a model of the sentence (1), for some $\xi < \beta$ there are at most γ_ξ ordinals in R_1 and hence fewer than α ordinals in R_1. On the other hand, since all the sentences (2) are in Σ, all the ordinals less than α are in R_1. This contradiction proves that Σ does not have a model.

Now suppose that Σ' is a subset of Σ of cardinal $< \alpha$. Then for some $\xi < \beta$, Σ is of cardinal $\leqslant \gamma_\xi$ and so there are at most γ_ξ of the sentences (2) in Σ'. It follows at once that Σ' has a model.

We have now shown that Σ is a locally consistent set of sentences with no model. Since Σ is a set of sentences of $L_{\alpha\alpha}$ of cardinal α this proves that the cardinal α is strongly incompact.

THEOREM 2.3. *Each successor cardinal is strongly incompact.*

PROOF. Suppose that α is a successor cardinal; say $\alpha=\beta^+$. Let Σ be the set of sentences of $L_{\alpha\alpha}$ consisting of σ_{WO}, the sentence

$$(\exists\langle v_\xi:\xi<\beta\rangle)(\forall v)\Big[\bigvee_{\xi<\beta}(v=v_\xi)\Big], \tag{1}$$

and, for each ordinal $\eta<\alpha$, the sentence

$$(\exists v)(Or_\eta(v)). \tag{2}$$

Suppose that Σ has a model, say $\langle\sigma,<\rangle$. Once again since Σ contains σ_{WO} we can assume that σ is an ordinal. Since $\langle\sigma,<\rangle$ is a model of all the sentences in (2), $\alpha\leqslant\sigma$; but by (1), σ is of cardinal at most β. This contradiction shows that Σ has no model.

On the other hand if Σ' is a subset of Σ of cardinal $<\alpha$ then we can show that Σ' has a model. For in this case there are at most β ordinals η such that the corresponding sentence (2) is in Σ'. Therefore if τ is the least ordinal greater than all these ordinals, $\text{card}(\tau)\leqslant\beta$ and hence $\langle\tau,<\rangle$ is a model of Σ'.

Σ is of cardinal α, and so we can again conclude that α is strongly incompact.

It is a consequence of the theorem we have just proved that every strongly accessible cardinal is strongly incompact and every accessible cardinal is incompact. Thus, assuming the G.C.H. every accessible cardinal is strongly incompact. If we do not make this assumption it is not known whether this still holds, but it should be noticed that HANF [1964] has shown that many accessible cardinals are strongly incompact, without using the G.C.H.

Theorem 2.1 implies that if α is a compact cardinal then α must be a strong limit cardinal, since if $\alpha\leqslant 2^\beta$, with $\beta<\alpha$, then α is 2^β-incompact and hence not compact. Theorem 2.2 shows that all weakly compact cardinals are regular and theorem 2.3 that all weakly compact cardinals are limit cardinals. Thus we can sum up the results of this section by saying that all compact cardinals are regular strong limit cardinals and all weakly compact cardinals are regular limit cardinals. Or, in other words, *compact cardinals are strongly inaccessible* and *weakly compact cardinals are inaccessible*. (Note that for the rest of this chapter, for brevity, we use 'inaccessible' in the sense of 'strongly inaccessible'.)

§3. Incompactness results for inaccessible cardinals

We have remarked that $\aleph_0$ is a compact cardinal, but apart from this example we have not yet shown that any compact cardinals exist. This is not surprising since the results of the previous section show that it is not possible to prove, within set theory, the existence of uncountable compact cardinals. We observed in this last section that compact cardinals are inaccessible and it is well known that the existence of uncountable inaccessibles cannot be proved within any standard system of set theory. This is because it can be proved within set theory, that if θ is an uncountable, inaccessible cardinal then the sets of rank less than θ form a model of the axioms of set theory. Therefore if we could prove the existence of uncountable inaccessibles we would have a proof within set theory of the existence of a model of set theory and thus of the consistency of set theory. This would contradict Gödel's second incompleteness theorem (GÖDEL [1931]). Since all compact cardinals other than $\aleph_0$ are uncountable inaccessibles, it follows that their existence also cannot be proved within any standard system of set theory.

However we now suppose that we have a (relatively consistent) strong axiom of infinity which guarantees the existence of inaccessibles and we suppose that

$$\theta_0,\theta_1,\ldots,\theta_\xi,\ldots$$

is an enumeration of the inaccessible cardinals in order of magnitude. (Thus we are assuming that our strong axiom of infinity implies the existence of an inaccessible θ_ξ for each ordinal ξ.) Given this assumption it makes sense to ask whether we can show that any of these inaccessibles is compact, or incompact. The result of Hanf, alluded to in the previous section, which we now prove, shows that, in a sense made precise in the statement of the theorem, most inaccessibles are strongly incompact.

THEOREM 3.1. HANF'S INCOMPACTNESS THEOREM
Each inaccessible $\alpha=\theta_\xi$, with $0<\xi<\alpha$ is strongly incompact. In particular, if $0<n<\omega$, then θ_n is strongly incompact.

PROOF. We consider the language $L_{\alpha\alpha}$ with the binary predicate letter P_0, two ternary predicate letters P_1 and P_2 and two unary predicate letters P_3 and P_4. We continue to write $x<y$ for $P_0(x,y)$. We let Σ be the set of sentences of $L_{\alpha\alpha}$ consisting of the sentence σ_{WO}, and the following:

(1). $(\forall tuvw)[P_1(t,u,v)\wedge P_1(t,u,w)\rightarrow v=w]$.

(2). $(\forall w)(\exists t)[t<w\wedge(\forall uv)[P_1(w,u,v)\rightarrow u<t\wedge v<w]]$.

(3). $(\forall wt)[P_3(w)\wedge t<w\rightarrow(\exists uv)[P_1(w,u,v)\wedge(t<v\vee t=v)]]$.

(4). $(\forall wuv)[u<w \wedge v<w \wedge (\forall t)[P_2(w,u,t) \leftrightarrow P_2(w,v,t)] \rightarrow u=v]$.

(5). $(\forall w)[P_4(w) \rightarrow (\exists t)[t<w \wedge (\forall uv)[P_2(w,u,v) \rightarrow v<t]]]$.

(It will emerge from the subsequent proof that $P_3(w)$ expresses the fact that w is a singular ordinal and $P_4(w)$ that w is not a strong limit cardinal.)

(6). $(\forall u)[Or_0(u) \vee P_3(u) \vee P_4(u) \vee \vee_{\eta<\xi} Or_{\theta_\eta}(u)]$.

(Thus this last sentence says that each ordinal is either 0, a singular ordinal, a non-strong limit cardinal or one of the inaccessibles $<\alpha$.)

(7). $(\exists u)(\forall v)[v<u \vee v=u]$.

(This sentence expresses the fact that there is a greatest ordinal in the model.)

(8). For each ordinal $\eta<\alpha$ the sentence $(\exists u)\, Or_\eta(u)$.

Σ is a set of sentences of $L_{\alpha\alpha}$ of cardinal α. Suppose that $\mathfrak{A}$ is a model of Σ. Since $\mathfrak{A}$ is a model of σ_{WO} we can assume that $\mathfrak{A}$ is of the form

$$\langle \sigma, <, R_1, R_2, R_3, R_4 \rangle,$$

where σ is an ordinal, $<$ is the 'less than' relation restricted to σ, R_1, R_2 are ternary relations on σ, and R_3, R_4 are subsets of σ.

We first show that R_3 is a set of singular ordinals. Suppose $\rho \in R_3$ and let

$$f_\rho = \{\langle \xi, \eta \rangle : \langle \rho, \xi, \eta \rangle \in R_1\}.$$

Since $\mathfrak{A}$ is a model of (1), f_ρ is a function. Since $\mathfrak{A}$ is a model of (2) there is an ordinal $\lambda<\rho$ such that

$$\text{domain}(f_\rho) \subseteq \lambda \quad \text{and} \quad \text{range}(f_\rho) \subseteq \rho$$

$\mathfrak{A}$ is a model of (3) and therefore the range of f_ρ is cofinal in ρ. It follows that ρ is a singular ordinal.

Next we show that the elements of R_4 are not strong limit cardinals. Suppose that ρ is a cardinal and $\rho \in R_4$. If $\eta<\rho$, let

$$g_\rho(\eta) = \{\xi : \langle \rho, \eta, \xi \rangle \in R_2\}.$$

Since $\mathfrak{A}$ is a model of (4), g_ρ is a one–one function defined on ρ. Since $\mathfrak{A}$ is a model of (5), there is some ordinal $\lambda<\rho$ such that the range of g_ρ is included in $\mathscr{P}(\lambda)$. It follows that $\rho \leqslant 2^{\text{card}(\lambda)}$, and thus that ρ is not a strong limit cardinal.

$\mathfrak{A}$ is a model of (6) and so, by the remarks above, each ordinal in σ is either 0, a singular ordinal, a non-strong limit cardinal or an inaccessible $<\alpha$. But since $\mathfrak{A}$ is a model of all the sentences (8), $\alpha \leqslant \sigma$, and because $\mathfrak{A}$ is a model of (7), σ is not a limit ordinal and so $\alpha<\sigma$. Thus the inaccessible α is in σ. We have therefore arrived at a contradiction. This contradiction shows that, contrary to our original hypothesis, Σ does not have a model.

On the other hand it is not difficult to see that Σ is locally consistent. For suppose Σ' is a subset of Σ of cardinal $<\alpha$. Let σ be the supremum of those ordinals η such that the formula $Or_\eta(u)$ occurs as a subformula of some sentence in Σ'. Since α is regular, we must have $\sigma<\alpha$. We construct the structure

$$\mathfrak{A}=\langle\sigma+1,<,R_1,R_2,R_3,R_4\rangle$$

as follows:

$<$ is the 'less than' relation restricted to $\sigma+1$. For each ordinal $\rho\leqslant\sigma$, if ρ is singular we choose the function f_ρ so that the domain of f_ρ is an ordinal less than ρ and the range of f_ρ is cofinal in ρ. Otherwise f_ρ is the empty function. We let

$$R_1=\{\langle\rho,\xi,\eta\rangle:\langle\xi,\eta\rangle\in f_\rho\}.$$

If $\rho\leqslant\sigma$ is a cardinal and for some $\tau<\rho$, $\rho\leqslant 2^\tau$, then we let g_ρ be a one–one map from ρ into $\mathscr{P}(\tau)$. Otherwise we let

$$g_\rho=\{\langle\eta,\{\eta\}\rangle:\eta<\rho\}.$$

We let

$$R_2=\{\langle\rho,\eta,\xi\rangle:\xi\in g_\rho(\eta)\}.$$

Finally we let R_3 be the set of all singular ordinals $\leqslant\sigma$ and we let R_4 be the set of all non-strong limit cardinals $\leqslant\sigma$. It is now easy to see that $\mathfrak{A}$ is a model of Σ'.

Thus we have shown that Σ is a locally consistent set of sentences of $L_{\alpha\alpha}$ of cardinal α with no model. It follows that α is α-incompact, and therefore strongly incompact. This completes the proof of Hanf's theorem.

§ 4. Measurable cardinals

We now discuss the relationship between the notion of a compact cardinal and that of a measurable cardinal which we introduced in ch. 6. Recall that a cardinal κ is said to be *measurable* if κ is uncountable and there is a non-principal ultrafilter on κ which is λ-complete for each $\lambda<\kappa$. In future we shall say that such an ultrafilter is *less than κ-complete*, which we shall write as *$<\kappa$-complete*.

Measurable cardinals play an important role in the model theory of infinitary languages because Łoš's theorem can be extended to ultraproducts of $L_{\kappa\kappa}$-structures whenever κ is measurable. We leave the proof of the following result to the reader; it is a simple extension of the corresponding result for first order languages.

THEOREM 4.1. ŁOŚ'S THEOREM FOR $L_{\kappa\kappa}$
Let $\{\mathfrak{A}_i: i \in I\}$ be a family of $L_{\kappa\kappa}$-structures, and let F be a $<\kappa$-complete ultrafilter on I. Then for each sequence $f = \langle f_\lambda: \lambda < \kappa \rangle \in (\prod_{i \in I} A_i)^\kappa$, and each formula ϕ of $L_{\kappa\kappa}$

$$\prod_{i \in I} \mathfrak{A}_i/F \vDash_{f/F} \phi \quad \textit{iff} \quad \{i \in I: \mathfrak{A}_i \vDash_{f(i)} \phi\} \in F.$$

As an immediate consequence we obtain

COROLLARY 4.2. *Each measurable cardinal is weakly compact.*

PROOF. Suppose κ is a measurable cardinal and that $\Sigma = \{\sigma_\lambda: \lambda < \kappa\}$ is a set of sentences of $L_{\kappa\kappa}$ of cardinal κ such that each subset of Σ of cardinal $<\kappa$ has a model. We want to show that Σ itself has a model. For each $\lambda < \kappa$ let $\mathfrak{A}_\lambda$ be a model of $\bigwedge_{\xi<\lambda} \sigma_\xi$ and let F be a $<\kappa$-complete non-principal ultrafilter on κ. We show that $\mathfrak{A} = \prod \mathfrak{A}_\lambda/F$ is the required model of Σ.

Since F is non-principal and $<\kappa$-complete, it contains the complements of all subsets of κ of cardinal $<\kappa$. Hence $\{\xi: \lambda < \xi < \kappa\} \in F$. But since

$$\{\xi: \lambda < \xi < \kappa\} \subseteq \{\xi < \kappa: \mathfrak{A}_\xi \vDash \sigma_\lambda\},$$

it follows that, for each $\lambda < \kappa$,

$$\{\xi < \kappa: \mathfrak{A}_\xi \vDash \sigma_\lambda\} \in F.$$

Therefore, by the previous theorem, $\mathfrak{A} \vDash \sigma_\lambda$ for each $\lambda < \kappa$. Thus $\mathfrak{A}$ is a model of Σ and our proof is finished.

It follows from this that measurable cardinals must be exceedingly large. In fact, in view of theorem 3.1, we have

COROLLARY 4.3. *Each inaccessible cardinal $\alpha = \theta_\xi$, with $0 < \xi < \alpha$ is non-measurable. Thus, in particular, $\theta_1, \theta_2, \ldots, \theta_\omega, \theta_{\omega+1}, \ldots, \theta_{\omega+\omega}, \ldots, \theta_{\omega_1}, \ldots$ are all non-measurable.*

Our next task will be to give a model theoretic characterization of measurable cardinals. Before embarking on this, however, we introduce some new definitions and notation.

Let κ be an infinite cardinal. Two $L_{\kappa\kappa}$-structures $\mathfrak{A}$ and $\mathfrak{B}$ are said to be $L_{\kappa\kappa}$-*equivalent* if each sentence σ of $L_{\kappa\kappa}$ which is true in $\mathfrak{A}$ is true also in $\mathfrak{B}$ and vice versa. $\mathfrak{A}$ is said to be an $L_{\kappa\kappa}$-extension of $\mathfrak{B}$, and $\mathfrak{B}$ is said to be an $L_{\kappa\kappa}$-*substructure* of $\mathfrak{A}$, if $\mathfrak{B} \subseteq \mathfrak{A}$ and for each formula ϕ of $L_{\kappa\kappa}$ and each sequence $b \in B^\kappa$,

$$\mathfrak{B} \vDash_b \phi \quad \text{iff} \quad \mathfrak{A} \vDash_b \phi.$$

If $\mathfrak{B}$ is an $L_{\kappa\kappa}$-substructure of $\mathfrak{A}$ we write $\mathfrak{B} \prec_\kappa \mathfrak{A}$. An extension $\mathfrak{A}$ of $\mathfrak{B}$ is said to be a *proper extension* if A properly contains B.

THEOREM 4.4. *Let κ be an uncountable cardinal, let $\mathfrak{A}_\kappa$ be the structure with domain κ whose relations consist of all those relations on κ of degree $<\kappa$, and let $\mathfrak{A}'_\kappa$ be the structure with domain κ whose relations consist of all unary relations on κ. Then the following conditions are equivalent:*

(a). *κ is measurable.*

(b). *$\mathfrak{A}_\kappa$ has a proper $L_{\kappa\kappa}$-extension.*

(c). *$\mathfrak{A}'_\kappa$ has a proper $L_{\kappa\kappa}$-extension.*

PROOF. (a)⇒(b). Assume (a); then there is a non-principal, $<\kappa$-complete ultrafilter F on κ. It is not difficult to see that $\mathfrak{A}_\kappa{}^\kappa/F$ is, up to isomorphism, a proper $L_{\kappa\kappa}$-extension of $\mathfrak{A}_\kappa$.

(b)⇒(c). This is trivial.

(c)⇒(a). Suppose (c) holds. Let

$$\mathfrak{B} = \langle B, \{S_\xi : \xi < 2^\kappa\}\rangle$$

be a proper $L_{\kappa\kappa}$-extension of

$$\mathfrak{A}'_\kappa = \langle \kappa, \{R_\xi : \xi < 2^\kappa\}\rangle .$$

Let b be an element of $B-\kappa$ and let $F = \{R_\xi : b \in S_\xi\}$. We have already proved (lemma 6.5.3) that F is a non-principal ultrafilter on κ. It remains only to show that if $\gamma < \kappa$ then F is γ-complete.

Suppose $\gamma < \kappa$ and $\{R_{\xi_\delta} : \delta < \gamma\}$ is a collection of elements of F. We show that $X = \bigcap_{\delta<\gamma} R_{\xi_\delta}$ is also an element of F.

There is some $\eta < 2^\kappa$ such that $X = R_\eta$. Hence the sentence

$$(\forall v)\Big[P_\eta(v) \leftrightarrow \bigwedge_{\delta<\gamma} P_{\xi_\delta}(v)\Big]$$

is true in $\mathfrak{A}'_\kappa$. It follows that it is also true in $\mathfrak{B}$ and so

$$S_\eta = \bigcap_{\delta<\gamma} S_{\xi_\delta} .$$

Therefore $b \in S_\eta$, and hence $X = R_\eta \in F$.

This shows that F is a $<\kappa$-complete, non-principal ultrafilter on κ, which is therefore measurable.

COROLLARY 4.5. *Each compact cardinal is measurable.*

PROOF. Let κ be a compact cardinal. We show that condition (b) of theorem 4.4 holds. It will follow at once that κ is measurable.

Let $L_{\kappa\kappa}$ have a predicate letter corresponding to each relation of degree

$<\kappa$ on κ. Let L′ be the language which is obtained from $L_{\kappa\kappa}$ by adding the constant symbols $\{c_\xi : \xi<\kappa\}$. Let Σ be the set of all sentences of L′ which hold in $(\mathfrak{A}_\kappa,\kappa)$, i.e. when for $\xi<\kappa$, the constant c_ξ is interpreted as the ordinal ξ.

Let L″ be the language which is obtained from L′ by adding a new constant symbol d and let

$$\Sigma''=\Sigma\cup\{c_\xi\neq d : \xi<\kappa\}.$$

We show that each subset of Σ'' of cardinal $<\kappa$ has a model. Let Σ_0'' be such a subset of Σ''. There is some $\xi<\kappa$ such that the sentence $c_\xi\neq d$ does not occur in Σ_0''. It follows that $(\mathfrak{A}_\kappa,\kappa,\xi)$ is a model of Σ_0''.

By hypothesis, κ is a compact cardinal, and each subset of Σ'' of cardinal $<\kappa$ has a model. Σ'' itself, therefore, has a model, say $(\mathfrak{B},x,x)$. Since $(\mathfrak{B},x,x)$ is a model of Σ, $\mathfrak{B}$ is, up to isomorphism, an $L_{\kappa\kappa}$-extension of $\mathfrak{A}_\kappa$. Since d has an interpretation in $\mathfrak{B}$ different from the interpretations of the c_ξ, $\mathfrak{B}$ is a proper extension of $\mathfrak{A}_\kappa$. This completes the proof.

§ 5. Keisler's ultrapower proof

We have already shown that none of the inaccessible cardinals θ_ξ, with $0<\xi<\theta_\xi$, is measurable. In this section we give a direct proof of this result, due to KEISLER [1964], using the ultrapower construction.

THEOREM 5.1. *If $0<\xi<\theta_\xi$ then θ_ξ is not measurable.*

PROOF. Suppose, on the contrary, that θ_ξ is a measurable cardinal and that F is a non-principal $<\theta_\xi$-complete ultrafilter on θ_ξ. Let $\mathfrak{A}=\langle\theta_\xi,<\rangle$, where $<$ is the usual order relation on the set of ordinals less than θ_ξ. Let $\mathfrak{A}'$ be the ultrapower

$$\mathfrak{A}'=\mathfrak{A}^{\theta_\xi}/F=\langle\theta_\xi{}^{\theta_\xi}/F,<_F\rangle.$$

F is ω-complete and therefore, by theorem 6.4.1, $<_F$ is a well-ordering. Hence there is a unique ordinal λ such that $\mathfrak{A}'$ is isomorphic to $\langle\lambda,<\rangle$. Let h be the isomorphism from $\mathfrak{A}'$ onto $\langle\lambda,<\rangle$.

We claim that $\theta_\xi<\lambda$. Given any two constant functions $f,g\in\theta_\xi{}^{\theta_\xi}$, if $f\neq g$, then $f/F\neq g/F$, because for each $\eta\in\theta_\xi$, $f(\eta)\neq g(\eta)$. The identity function $id\in\theta_\xi{}^{\theta_\xi}$ exceeds any constant function except on a set containing fewer than θ_ξ elements. F is non-principal and $<\theta_\xi$-complete. Therefore no subset of θ_ξ of cardinal $<\theta_\xi$ is in F. Hence, by Łoš's theorem, $f/F<_F id/F$, for each constant function f. Thus $h(id/F)$ has at least θ_ξ predecessors. It follows at once that $\theta_\xi<\lambda$ and we can deduce from this that there is some $g\in\theta_\xi{}^{\theta_\xi}$ such that $\theta_\xi=h(g/F)$.

We define the sets X_i, $0 \leqslant i \leqslant 3$, and Y_ζ, $\zeta < \xi$, as follows:

$$X_0 = \{\eta < \theta_\xi : g(\eta) = 0\},$$
$$X_1 = \{\eta < \theta_\xi : g(\eta) \text{ is not a cardinal}\},$$
$$X_2 = \{\eta < \theta_\xi : g(\eta) \text{ is a singular cardinal}\},$$
$$X_3 = \{\eta < \theta_\xi : g(\eta) \text{ is a regular accessible cardinal}\},$$

and, for $\zeta < \xi$,

$$Y_\zeta = \{\eta < \theta_\xi : g(\eta) = \theta_\zeta\}.$$

These sets are pairwise disjoint. Therefore, by lemma 6.1.12, precisely one of them is in the ultrafilter F. We show, by considering each possibility in turn, that this leads to a contradiction.

If $X_0 \in F$, then, by Łoš's theorem, $\theta_\xi = h(g/F) = 0$, which is obviously false.

Before we consider the next case we observe that for each cardinal $\alpha < \theta_\xi$, the set

$$\{f/F \in \theta_\xi^{\theta_\xi}/F : f/F <_F \alpha^*/F\}$$

is of cardinal α, where $\alpha^* \in \theta_\xi^{\theta_\xi}$ is the function given by $\alpha^*(\eta) = \alpha$, for each $\eta < \theta_\xi$. For if $f/F <_F \alpha^*/F$, then

$$Z_f = \{\eta < \theta_\xi : f(\eta) < \alpha\} \in F.$$

But

$$Z_f = \bigcup_{\zeta < \alpha} \{\eta < \theta_\xi : f(\eta) = \zeta\},$$

and so it follows from lemma 6.1.12 and the fact that F is $< \theta_\xi$-complete, that for some $\zeta < \alpha$,

$$\{\eta < \theta_\xi : f(\eta) = \zeta\} \in F.$$

Hence $f/F = \zeta^*/F$. Since $\text{card}(\{\zeta^*/F : \zeta < \alpha\}) = \alpha$, the result follows at once.

This remark shows that if $Y_\zeta \in F$ for some $\zeta < \xi$ then $g/F = \theta_\zeta^*/F$, and so $\theta_\xi = h(g/F)$ would have only θ_ζ predecessors and so would be identical with θ_ζ, an absurdity.

Suppose now that $X_1 \cup X_2 \in F$. For each ordinal $\gamma < \theta_\xi$ which is either not a cardinal or is a singular cardinal, there is some $\mu_\gamma < \gamma$ (namely $\text{cf}(\gamma)$) and a function G_γ mapping μ_γ onto a cofinal subset of γ. Let $\mu = h(m/F)$ where $m \in \theta_\xi^{\theta_\xi}$ is such that for $\eta \in X_1 \cup X_2$,

$$m(\eta) = \mu_{g(\eta)}.$$

Since $X_1 \cup X_2 \in F$, it is easy to see that $\mu < \theta_\xi$. Let G be the function mapping μ into θ_ξ such that, whenever $h(f/F) < \mu$, we have $G(h(f/F)) = h(j/F)$,

where j is such that $j(\eta)=G_{g(\eta)}(f(\eta))$ for $\eta\in X_1\cup X_2$ with $f(\eta)<\mu_{g(\eta)}$. (It is easy to see that G is well defined.)

Now for each $\alpha<\theta_\xi$ there is some $a\in\theta_\xi^{\theta_\xi}$ such that $\alpha=h(a/F)$; using the axiom of choice we can construct a function b such that $b(\eta)<\mu_{g(\eta)}$ and $a(\eta)\leqslant G_{g(\eta)}(b(\eta))$ whenever $\eta\in X_1\cup X_2$ and $a(\eta)<g(\eta)$.

Let $\beta=h(b/F)$; it follows immediately from the properties of b that $\beta<\mu$ and $\alpha\leqslant G(\beta)$. Thus for each $\alpha<\theta_\xi$ there is a $\beta<\mu$ such that $\alpha\leqslant G(\beta)$, which implies that θ_ξ is not a regular cardinal, which is impossible.

The final case we have to consider is $X_3\in F$. For each regular accessible cardinal $\gamma<\theta_\xi$ we can find some $\sigma_\gamma<\gamma$ such that $\gamma\leqslant 2^{\sigma_\gamma}$. Thus for each $\alpha<\gamma$ we can find a subset $H_\gamma(\alpha)$ of σ_γ such that if $\alpha,\beta<\gamma$ and $\alpha\neq\beta$, then $H_\gamma(\alpha)\neq H_\gamma(\beta)$. Let s be a function such that $s(\eta)=\sigma_{g(\eta)}$ for all $\eta\in X_3$ and let $\sigma=h(s/F)$. Since $X_3\in F$, $\sigma<\theta_\xi$.

Now for each $h(f/F)<\theta_\xi$ define $H(h(f/F))$ to be the set of all $h(j/F)<\lambda$ such that $j(\eta)\in H_{g(\eta)}(f(\eta))$ whenever $f(\eta)<g(\eta)$ and $\eta\in X_3$. It is easily seen that H is well-defined. For each f such that $h(f/F)<\theta_\xi$ we have

$$\{\eta<\theta_\xi: f(\eta)<g(\eta)\}\in F,$$

so $H(h(f/F))\subseteq\sigma$. If $\alpha,\beta<\theta_\xi$ and $\alpha\neq\beta$, then $H(\alpha)\neq H(\beta)$; for if $\alpha=h(a/F)$ and $\beta=h(b/F)$, and f is a function such that $f(\eta)$ is a member of $H_{g(\eta)}(a(\eta))$ or $H_{g(\eta)}(b(\eta))$ but not both whenever $\eta\in X_3$, $a(\eta)<g(\eta)$, $b(\eta)<g(\eta)$ and $a(\eta)\neq b(\eta)$, then $h(f/F)$ is in $H(\alpha)$ or $H(\beta)$ but not both.

Thus H is a one-one map of θ_ξ into the set of all subsets of σ. Hence we have

$$\sigma<\theta_\xi\leqslant 2^\sigma,$$

which contradicts that θ_ξ is inaccessible.

We have now arrived at a contradiction in each case, and the proof is complete.

§ 6. Measurable cardinals and the axiom of constructibility

In this section we explain the proof due to SCOTT [1961] that the existence of measurable cardinals contradicts the axiom of constructibility, that is $V=L$, in the sense of GÖDEL [1940].

We begin with some definitions and remarks. We shall distinguish between *sets* and *classes*, a set being a class which is a member of another class. We shall reserve upper case letters for classes and lower case letters for sets. The class of all sets will be denoted by V. For each class X, we shall

denote the class of all subsets of X by $\mathscr{P}(X)$. A class X is said to be *transitive* if $X \subseteq \mathscr{P}(X)$, or in other words, if whenever $x \in X$ and $y \in x$, then $y \in X$.

If x,y are sets we let

$$x - y = \{z : z \in x \text{ and } z \notin y\},$$
$$\bigcup x = \{z : \text{for some } y \in x,\ z \in y\},$$
$$\bar{x} = \{\langle y,z\rangle : \langle z,y\rangle \in x\},$$
$$x \mid y = \{\langle z,w\rangle : \text{for some } u, \langle z,u\rangle \in x \text{ and } \langle u,w\rangle \in y\},$$

and

$$E \restriction x = \{\langle y,z\rangle : y \in z \text{ and } z \in x\}.$$

For our purposes we will take the *axiom of constructibility* to be the following:

If M is a class such that

(1). $M \subseteq \mathscr{P}(M) \subseteq \bigcup_{x \in M} \mathscr{P}(x)$,

and

(2). $x-y$, $\bigcup x$, $\bar{x}$, $x|x$ and $E \restriction x$ are in M for all $x,y \in M$, then $V = M$.

The proof of Scott's theorem may be outlined as follows: Assume both the axiom of constructibility and that a measurable cardinal exists. Let κ be the smallest measurable cardinal and suppose that F is a non-principal, $<\kappa$-complete ultrafilter on κ. One can then show that the 'ultrapower' V^κ/F is isomorphic to V. (Of course really V^κ/F is a fiction since V is a proper class; what we actually use is a genuine class V^* which behaves in the same way as V^κ/F would, if it existed.) An application of Łoš's theorem immediately yields the result that $\kappa = \kappa^*$, where κ^* is the member of V^κ/F corresponding to the function with constant value κ. But on the other hand, since F is non-principal, we must have $\kappa < \kappa^*$ and this is the required contradiction.

Our proof will take the form of a series of lemmas, culminating in a contradiction. Assume once and for all that the axiom of constructibility holds and that measurable cardinals exist. Let κ be the smallest measurable cardinal and let F be a non-principal, $<\kappa$-complete ultrafilter on κ. Define

$$Q = \{\langle f,g\rangle : f,g \in V^\kappa \text{ and } \{\xi < \kappa : f(\xi) = g(\xi)\} \in F\},$$

and

$$E = \{\langle f,g\rangle : f,g \in V^\kappa \text{ and } \{\xi < \kappa : f(\xi) \in g(\xi)\} \in F\}.$$

As in the proof of lemmas 5.1.1 and 5.1.3, it is easy to show that Q is an equivalence relation and is a congruence relation with respect to E. And since F is countably complete, it is clear that E is a well-founded relation on V^κ. We now prove

LEMMA 6.1. (i). *If* $\{h\in V^\kappa: hEf\}=\{h\in V^\kappa: hEg\}$, *then* fQg.

(ii). $\{h\in V^\kappa: hEf\}=\{h\in V^\kappa$: *for some* k, $k\in(\bigcup_{\xi<\kappa} f(\xi)\cup\{0\})^\kappa$, kEf *and* $hQk\}$.

PROOF. (i). Suppose not fQg; then $\{\xi: f(\xi)\neq g(\xi)\}\in F$. Using the axiom of choice we can find a function $h\in V^\kappa$ such that either

$$\{\xi: h(\xi)\in f(\xi)\}\in F \quad \text{and} \quad \{\xi: h(\xi)\in g(\xi)\}\notin F,$$

or

$$\{\xi: h(\xi)\in f(\xi)\}\notin F \quad \text{and} \quad \{\xi: h(\xi)\in g(\xi)\}\in F.$$

In either case we have

$$\{k\in V^\kappa: kEf\}\neq\{k\in V^\kappa: kEg\}.$$

(ii). Let

$$A=\{h\in V^\kappa: \text{for some } k\in(\bigcup_{\xi<\kappa} f(\xi)\cup\{0\})^\kappa, kEf \text{ and } hQk\}.$$

Clearly $A\subseteq\{h\in V^\kappa: hEf\}$. Conversely if $h\in V^\kappa$ and hEf, then

$$X=\{\xi<\kappa: h(\xi)\in f(\xi)\}\in F.$$

Define

$$k\in(\bigcup_{\xi<\kappa} f(\xi)\cup\{0\})^\kappa$$

by putting $k(\xi)=h(\xi)$ for $\xi\in X$ and $k(\xi)=0$ otherwise. Then plainly kEf and hQk, and the reverse inclusion follows.

A well-known consequence of the axiom of constructibility is that one can define a function ρ from sets to ordinals such that if $x\in y$ then $\rho(x)<\rho(y)$ and for each ordinal λ, the class $\{x:\rho(x)<\lambda\}$ is a set. For each $f\in V^\kappa$ we define $\lambda(f)$ to be the least ordinal μ such that there is some $g\in V^\kappa$ with gQf and $\rho(g)=\mu$. Also we let

$$f_Q=\{g: fQg \quad \text{and} \quad \rho(g)=\lambda(f)\}.$$

Clearly

$$f_Q=g_Q \quad \text{iff} \quad fQg.$$

Let

$$V_Q=\{f_Q: f\in V^\kappa\},$$

and let $\in_Q$ be the relation on V_Q given by

$$f_Q\in_Q g_Q \quad \text{iff} \quad fEg.$$

Since E is a congruence relation with respect to Q, this relation is well-defined. Since E is a well-founded relation, so also is $\in_Q$. In addition it is

easy to verify that the structure

$$\langle V_Q, \in_Q \rangle$$

is a model of the axiom of extensionality. That is, if $f_Q \neq g_Q$, then there is some $h_Q \in V_Q$ which is in the relation $\in_Q$ to one of f_Q, g_Q, but not to the other.

LEMMA 6.2. *For each $f \in V$, $\{g_Q : g_Q \in_Q f_Q\}$ is a set.*

PROOF. By lemma 6.1,

$$\{g_Q : g_Q \in_Q f_Q\} = \{k_Q : k \in (\bigcup_{\xi<\kappa} f(\xi) \cup \{0\})^\kappa \text{ and } k_Q \in_Q f_Q\},$$

which is certainly a set.

LEMMA 6.3. *$\langle V_Q, \in_Q \rangle$ is isomorphic to a unique transitive structure $\langle M, \in \rangle$.*

PROOF. In view of the previous lemma and the fact that the relation $\in_Q$ is well-founded we can define by recursion the function T such that for all $y \in V_Q$

$$T(y) = \{T(x) : x \in V_Q \text{ and } x \in_Q y\}.$$

Let M be the range of T. It is easy to see that M is a transitive class, and that T is an isomorphism of $\langle V_Q, \in_Q \rangle$ onto $\langle M, \in \rangle$.

LEMMA 6.4. $M \subseteq \mathscr{P}(M) \subseteq \bigcup_{x \in M} \mathscr{P}(x)$.

PROOF. Since $\langle M, \in \rangle$ is a transitive structure, the first inclusion holds. To prove the second assume $y \in \mathscr{P}(M)$. Using the axiom of choice we can find $z \in \mathscr{P}(V^\kappa)$ such that

$$y = \{T(g_Q) : g \in z\}.$$

Let $f \in V^\kappa$, be defined by

$$f(\xi) = \{g(\xi) : g \in z\}, \quad \text{for each } \xi < \kappa.$$

Then clearly $y \subseteq T(f_Q)$, and the second inclusion follows.

For each formula $\Phi(v_0, \ldots, v_n)$ of set theory without bound class variables we let $\Phi^{(M)}(v_0, \ldots, v_n)$ be the formula obtained by relativising all the quantifiers to M. Thus, intuitively, for $x_0, \ldots, x_n \in M$, the formula $\Phi^{(M)}(x_0, \ldots, x_n)$ says that the formula Φ is satisfied by $x_0, \ldots, x_n$ in the structure $\langle M, \in \rangle$.

Since $\langle M, \in \rangle$ is isomorphic to an 'ultrapower' of V, one can prove by induction the following special version of Łoś's theorem. We leave the proof to the reader.

LEMMA 6.5. *If $f_0, \ldots, f_n \in V^\kappa$, then for each Φ*

$$\Phi^{(M)}(T(f_{0Q}), \ldots, T(f_{nQ})) \quad \textit{iff} \quad \{\xi < \kappa : \Phi(f_0(\xi), \ldots, f_n(\xi))\} \in F.$$

LEMMA 6.6. *The class M satisfies condition* (2) *of the axiom of constructibility, and therefore* $V = M$.

PROOF. This is an almost immediate consequence of the previous lemma. To illustrate how the proof goes, we show that M is closed under the union operation; the other cases are dealt with similarly.

Let $\Phi(v_0,v_1)$ be the formula

$$(\forall v_2)[v_2 \in v_1 \leftrightarrow (\exists v_3)[v_3 \in v_0 \wedge v_2 \in v_3]]$$

thus $\Phi(x,y)$ iff $y = \bigcup x$.

The sentence $(\forall v_0)(\exists v_1)\,\Phi(v_0,v_1)$ holds in V and hence, by lemma 6.5,

$$(\forall v_0)(\exists v_1)\Phi(v_0,v_1)^{(M)} = (\forall v_0)[v_0 \in M \rightarrow (\exists v_1)[v_1 \in M \wedge \Phi^{(M)}(v_0,v_1)]]$$

holds in M. But since M is transitive, for $x,y \in M$,

$$\Phi(x,y) \quad \text{iff} \quad \Phi^{(M)}(x,y).$$

Hence if $x \in M$, there is some $y \in M$ such that $\Phi(x,y)$, that is, such that $y = \bigcup x$. Thus $\bigcup x \in M$, and so M is closed under the union operation.

For each $x \in V$, let x^+ be the constant function in V with value x, and let $x^* = T(x^+_Q)$.

LEMMA 6.7. *If* $x_0,\ldots,x_n \in V$ *then*

$$\Phi(x_0{}^*,\ldots,x_n{}^*) \quad \textit{iff} \quad \Phi(x_0,\ldots,x_n).$$

In particular, if $\alpha \in V$ *is an ordinal, then so also is* α^*.

PROOF. The first part is an immediate consequence of the previous lemmas, the second of the fact that there is a formula of set theory characterizing ordinals.

If we apply this lemma to the formula which says that κ is the smallest measurable cardinal, we can conclude that κ^* is also the smallest measurable cardinal and so $\kappa = \kappa^*$.

Let $id \in \kappa^\kappa$ be the identity function, that is

$$id = \{\langle \xi,\xi \rangle : \xi \in \kappa\},$$

and let

$$\mu = T(id_Q).$$

By lemma 6.6, μ is an ordinal. Now we have

LEMMA 6.8. *If* $\lambda < \kappa$ *then* $\lambda^* < \mu < \kappa^*$.

PROOF. Suppose $\lambda < \kappa$. Then $\{\xi : \lambda < \xi < \kappa\} \in F$ and hence $\lambda^+ E id$, and so $T(\lambda^+_Q) \in T(id_Q)$, that is $\lambda^* < \mu$. Similarly $\mu < \kappa^*$.

By lemma 6.5, the map which sends the element $x \in V$ onto x^* is one–one. Therefore $\{\lambda^* : \lambda < \kappa\}$ is of cardinal κ. Thus, by lemma 6.8, μ is of cardinality at least κ. But this contradicts the fact that $\kappa = \kappa^*$ and $\mu < \kappa^*$. We have therefore at last arrived at the contradiction we promised. Thus we have proved

THEOREM 6.9. SCOTT'S THEOREM
The existence of measurable cardinals is inconsistent with the axiom of constructibility.

§7. Historical and bibliographical remarks

Infinitary predicate languages were first formulated explicitly in TARSKI [1958]. The problem of axiomatizing these languages was discussed by Carol Karp in her unpublished doctoral dissertation; a systematic exposition of her results is given in KARP [1964].

The results of §§4 and 5 were first obtained by HANF [1964], where he shows that the members of a very extensive class M of inaccessible cardinals are all strongly incompact. The class M includes all inaccessibles θ_ξ with $0 < \xi < \theta_\xi$ (theorem 3.1); if $\theta'_0, \theta'_1, \ldots$ is an enumeration of the 'fixed points' of the sequence $\theta_0, \theta_1, \ldots$ (i.e. the θ_ξ for which $\xi = \theta_\xi$), then M also includes all θ'_ξ for which $\xi < \theta'_\xi$ etc. Hanf shows, in fact, that M contains all inaccessibles which can, roughly speaking, be described in terms of smaller ordinals. The proof of theorem 3.1 already shows how one would go about establishing such an assertion: if $\Phi(x)$ is the matrix of sentence (6) of §3 then the statement $(\forall x)\, \Phi(x)$ holds in an ordinal structure $\langle \eta, <, \ldots \rangle$ iff $\eta \leqslant \theta_\xi$. Thus θ_ξ is the largest ordinal η for which $\langle \eta, <, \ldots \rangle \models (\forall x)\, \Phi(x)$. This notion of *describability* is explicitly formulated in HANF and SCOTT [1961]; a more detailed exposition is given in LEVY [1967].

The concept of a measurable cardinal stems back, in principle, to ULAM [1930]. TARSKI [1939] showed that each measurable cardinal is inaccessible. In fact very little else was known about measurable cardinals until 1960 when Tarski used Hanf's results to show that the first measurable cardinal, if it exists, must be of gargantuan proportions (TARSKI [1962]). A consequence of his discovery is that, if it is consistent (with respect to the ordinary axioms of set theory) to assume the existence of all the $\theta_\xi, \theta'_\xi, \theta''_\xi, \ldots$ etc., then one can add the premise that they are all non-measurable without destroying that consistency.

The result of SCOTT [1961] that the existence of measurable cardinals

contradicts the axiom of constructibility provided the first indication that the presence of measurable cardinals affects not only the 'length' of the universe of sets, i.e. the number of strong axioms of infinity satisfied, but also, so to speak, its 'width'. For Scott's result can be extended to show that if F is a $<\kappa$-complete, non-principal ultrafilter on an uncountable cardinal κ, then F cannot be constructible. It follows that if $\mathfrak{A}$ is a model of set theory together with the axiom that an uncountable measurable cardinal exists, then the submodel $\mathfrak{A}_L$ of sets constructible in $\mathfrak{A}$ cannot contain the non-principal, $<\kappa$-complete ultrafilter F on the first measurable cardinal in $\mathfrak{A}$. In some sense, therefore, $\mathfrak{A}_L$ is 'thinner' than $\mathfrak{A}$. On the other hand $\mathfrak{A}_L$ and $\mathfrak{A}$ may be regarded as being of equal 'length' in view of the fact that $\mathfrak{A}_L$ contains all the ordinals of $\mathfrak{A}$. Thus assuming the existence of a measurable cardinal provides a much more powerful and risky axiom of infinity than that obtained by assuming the existence of the $\theta_\xi, \theta'_\xi, \theta''_\xi$ etc.

The surprising connection between the width of the constructible universe and the presence of measurable cardinals established by Scott was further investigated by Rowbottom and later by Gaifman. In his unpublished doctoral dissertation Rowbottom shows that, if a measurable cardinal exists, then the constructible continuum is countable, i.e. there are only countably many constructible real numbers. (This is a profound and arresting instance of the so-called 'Skolem paradox', in which a set uncountable in one model may yet be countable in a larger model.) GAIFMAN [1964] extends this result and proves that under the same assumption, if α is an ordinal definable by some formula ϕ of set theory, then the ordinal $\alpha^{(L)}$ is countable, where $\alpha^{(L)}$ is the constructible ordinal satisfying the relativization $\phi^{(L)}$ of the formula ϕ to the constructible universe. Thus $\aleph_1^{(L)}, \aleph_2^{(L)}, \ldots, \aleph_\omega^{(L)}$ etc., are all countable.

The most recent discoveries in this field are those of Silver and Solovay. In order to state them we require a few definitions. Let us say that a second order formula is Σ_n^1 (Π_n^1) if it is in prenex normal form, all the second order quantifiers appear at the beginning, and there are $n-1$ alternations of these of which the first is existential (universal). For example,

$$(\forall f)(\exists g)(\forall x)[f(x) < g(x)]$$

is a Π_2^1-formula. A set of natural numbers is Σ_n^1 (Π_n^1) if it is definable in $\langle \omega, < \rangle$ by a Σ_n^1- (Π_n^1-) formula. It is said to be Δ_n^1 if it is both Σ_n^1 and Π_n^1.

Now SILVER [1966] and SOLOVAY [1967] have shown that, if a measurable cardinal exists, then the non-constructible set of natural numbers, shown by Rowbottom to exist, may be taken to be Δ_3^1. This should be contrasted with

a theorem of SHOENFIELD [1961] to the effect that every Σ^1_2-set of natural numbers is constructible. It is known that the existence of a non-constructible Δ^1_3-set of natural numbers is consistent with the usual axioms for set theory (see MATHIAS [1967]).

Finally, a brief summary of the known results and unresolved problems about the relative sizes of compact and measurable cardinals. We know that every measurable cardinal is weakly compact; HANF and SCOTT [1961] have shown that the first measurable cardinal is strictly larger than the first uncountable weakly compact inaccessible, if they exist. Whether the first uncountable *compact* cardinal coincides with the first measurable cardinal is still unknown. But Příkrý (in an unpublished paper) has shown that, if one can prove in set theory that every uncountable compact cardinal is strictly greater than the smallest measurable cardinal, then one can also prove in set theory that each compact cardinal $\lambda > \aleph_0$ is the λ-th measurable cardinal. Thus, under these conditions, each uncountable compact cardinal is a 'fixed point' of the sequence of measurable cardinals, just as each measurable cardinal is a 'fixed point' of the sequence of inaccessibles.

SOME SUGGESTIONS FOR FURTHER READING

A complete bibliography of model theory up to the end of 1963 is given in ADDISON, HENKIN and TARSKI [1965]. This volume contains an interesting and varied selection of recent contributions to model theory. Publications in this field are reviewed in *Mathematical Reviews* and *The Journal of Symbolic Logic.*

We have covered in this book most of the applications of ultraproducts to model theory, but to prevent the book from becoming too large we have not included them all. The most notable omissions are the work of AX and KOCHEN [1965], [1965i] and [1966], on p-adic fields and Keisler's work on limit ultrapowers [1963] and limit ultraproducts [1965]. Two useful surveys of the various applications of ultraproducts are KEISLER [1965i] and CHANG [1967].

In order to find out about those parts of model theory not treated in this volume, the reader should consult the books of CHANG and KEISLER [1966], KREISEL and KRIVINE [1967] and ROBINSON [1963], and also the recent lecture notes of MORLEY [1968].

In the bibliography that follows we have only given references to works cited in this book.

BIBLIOGRAPHY

ADDISON, J., L. HENKIN and A. TARSKI (Editors)

[1965] The theory of models. *Proceedings of the 1963 international symposium at Berkeley*, North-Holland, Amsterdam.

AX, J., and S. KOCHEN

[1965] Diophantine problems over local fields I. *Am. J. Math.*, **87**, 605–630.

[1965i] Diophantine problems over local fields II: a complete set of axioms for p-adic number theory. *Am. J. Math.*, **87**, 631–648.

[1966] Diophantine problems over local fields III: decidable fields. *Ann. Math.*, **83**, 437–456.

BACHMANN, H.

[1967] *Transfinite Zahlen*, 2nd ed., Springer, Berlin.

BAUER-MENGELBERG, S.

[1965] Review of GÖDEL [1962]. *J. Symb. Logic*, **30**, 359–362.

[1966] Review of the translation of GÖDEL [1931] in DAVIS [1965]. *J. Symb. Logic*, **31**, 486–489.

BETH, E. W.

[1953] On Padoa's method in the theory of definition. *Indag. Math.*, **15**, 330–339.

BIRKHOFF, G.

[1967] *Lattice theory*, 3rd ed., Am. Math. Soc. *Coll. Publ.*, vol. XXV.

BOOLE, G.

[1847] *The mathematical analysis of logic*, MacMillan, Barclay and MacMillan, Cambridge. (Reprinted by Basil Blackwell, Oxford, 1947).

CANTOR, G.

[1895] Beiträge zur Begründung der transfiniten Mengenlehre. *Math. Ann.*, **96**, 481–512. (English translation in CANTOR [1915]).

[1915] *Contributions to the founding of the theory of transfinite numbers*, translated and provided with an introduction and notes by Philip E. B. Jourdain (Reprinted by Dover Publications, New York n.d.)

CARTAN, H.

[1937] Théorie des filtres. *Compt. rend. hebd. séances Acad. Sci.*, **205**, 595–598.

[1937i] Filtres et ultrafiltres. *Ibid.*, 777–779.

CHANG, C. C.

[1965] A note on the two-cardinal problem. *Proc. Am. Math. Soc.*, **16**, 1148–1155.

[1965i] A simple proof of the Rabin–Keisler theorem. *Bull. Am. Math. Soc.*, **71**, 642–643.

[1967] Ultraproducts and other methods of constructing models. *Sets, models and recursion theory*, edited by J. N. Crossley, North-Holland, Amsterdam, 85–121.

CHANG, C. C. and H. J. KEISLER

[1962] Applications of ultraproducts of pairs of cardinals to the theory of models. *Pacific J. Math.*, **12**, 835–845.

[1966] *Continuous model theory*, Princeton U.P.

CHURCH, A.

[1936] A note on the Entscheidungsproblem. *J. Symb. Logic*, **1**, 40–41.

[1956] *Introduction to mathematical logic*, Princeton U.P.

CRAIG, W.
[1957] Three uses of the Herbrand–Gentzen theorem in relating model theory to proof theory. *J. Symb. Logic*, **22**, 269–285.

DAVIS, M. (Editor)
[1965] *The undecidable; basic papers on undecidable propositions, unsolvable problems and computable functions*, Raven Press, New York.

DREBEN, B., and S. AANDERAA
[1964] Herbrand analyzing functions. *Bull. Am. Math. Soc.*, **70**, 697–698.

DREBEN, B., P. ANDREWS and S. AANDERAA
[1963] False lemmas in Herbrand. *Bull. Am. Math. Soc.*, **69**, 699–706.

DREBEN, B., and J. DENTON
[1966] A supplement to Herbrand. *J. Symb. Logic*, **31**, 393–398.

EASTON, W.
[1964] *Powers of regular cardinals*. Doctoral dissertation, Princeton University.

ERDOS, P., A. HAJNAL and R. RADO
[1965] Partition relations for cardinal numbers. *Acta Math.*, **16**, 93–196.

FEFERMAN, S.
[1952] Review of RASIOWA and SIKORSKI [1951]. *J. Symb. Logic*, **17**, 72.

FRAENKEL, A.
[1966] *Abstract set theory*, 2nd ed., North-Holland, Amsterdam.

FRAÏSSÉ, R.
[1954] Sur l'extension aux relations de quelques propriétés des ordres. *Ann. Sci. École Norm. Sup.*, **71**, 363–388.

FRAYNE, T., A. MOREL and D. SCOTT
[1962] Reduced direct products. *Fund. Math.*, **51**, 195–228.

FUHRKEN, G.
[1964] Skolem-type normal forms for first order languages with a generalized quantifier. *Fund. Math.*, **54**, 291–302.
[1965] Languages with the added quantifier 'there exist at least $\aleph_\alpha$'. *The theory of models*, edited by J. Addison, L. Henkin and A. Tarski, North-Holland, Amsterdam, 121–131.

GAIFMAN, H.
[1964] Measurable cardinals and constructible sets. *Notices Am. Math. Soc.*, **11**, 771.

GILBERT, M.
[1965] *Plough my own furrow*, Longmans, London.

GILLMAN, L., and M. JERISON
[1960] *Rings of continuous functions*, Van Nostrand, Princeton.

GÖDEL, K.
[1930] Die Vollständigkeit der Axiome des logischen Funktionenkalküls. *Monatsh. Math. Phys.*, **37**, 349–360. (English translation in VAN HEIJENOORT [1967], 582–591.)
[1931] Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Ibid.*, **38**, 173–198. (English translations in GÖDEL [1962], DAVIS [1965], 4–38 and VAN HEIJENOORT [1967], 596–617. The first two of these are not altogether reliable (see BAUER-MENGELBERG [1965] and [1966]).)

[1940] *The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory*, Princeton U.P.

[1962] *On formally undecidable propositions of Principia Mathematica and related systems*, translated by B. Meltzer, with introduction by R. B. Braithwaite, Oliver and Boyd, Edinburgh.

GOODSTEIN, R.

[1963] *Boolean algebra*, Pergamon Press, Oxford.

HALMOS, P.

[1963] *Lectures on Boolean algebras*, Van Nostrand, Princeton.

HALMOS, P., and H. VAUGHAN

[1950] The marriage problem. *Am. J. Math.*, **72**, 214–215.

HANF, W.

[1960] Models of languages with infinitely long expressions. *International congress for logic, methodology and philosophy of science, abstracts of contributed papers*, Stanford, California, 24.

[1964] Incompactness in languages with infinitely long expressions, *Fund. Math.*, **53**, 309–324.

HANF, W., and D. SCOTT

[1961] Classifying inaccessible cardinals. *Notices Am. Math. Soc.*, **8**, 445.

HAUSDORFF, F.

[1914] *Grundzüge der Mengenlehre*, Leipzig. (Reprinted by the Chelsea Publishing Co., New York, 1949.)

HEIJENOORT, J. VAN

[1967] *From Frege to Gödel, a source book in mathematical logic 1879–1931*, Harvard U.P.

HENKIN, L.

[1949] The completeness of the first order functional calculus, *J. Symb. Logic*, **14**, 159–166.

HERBRAND, J.

[1928] Sur la théorie de la démonstration. *Compt. rend. hebd. séances Acad. Sci.*, **126**, 1274–1276.

[1930] Recherches sur la théorie de la démonstration, *Trav. Soc. Sci. Lett. Varsovie, cl. III*, **33**.

HEWITT, E.

[1948] Rings of real-valued continuous functions. *Trans. Am. Math. Soc.*, **64**, 45–99.

HIGMAN, G.

[1963] *Formal arithmetic*, Mathematical Institute, Oxford.

HILBERT, D., and P. BERNAYS

[1934] *Grundlagen der Mathematik I*, Springer, Berlin.

HUNTINGTON, E. V.

[1904] Sets of independent postulates for the algebra of logic. *Trans. Am. Math. Soc.*, **5**, 288–309.

JANICZAK, A.

[1950] A remark concerning the decidability of complete theories. *J. Symb. Logic*, **15**, 277–279.

JÓNSSON, B.
[1956] Universal relational systems. *Math. Scand.*, **4**, 193–208.
[1960] Homogeneous universal relational systems. *Math. Scand.*, **8**, 137–142.
KALMÁR, L.
[1934–35] Über die axiomatisierbarkeit des Aussagenkalküls. *Acta Sci. Math. (Szeged)*, **7**, 222–243.
KAPLANSKY, I.
[1954] *Infinite Abelian groups*, the University of Michigan Press, Ann Arbor.
KARP, C. R.
[1964] *Languages with expressions of infinite length*, North-Holland, Amsterdam.
KEISLER, H. J.
[1961] Ultraproducts and elementary classes. *Indag. Math.*, **23**, 477–495.
[1963] Limit ultrapowers. *Trans. Am. Math. Soc.*, **107**, 382–408.
[1964] On cardinalities of ultraproducts. *Bull. Am. Math. Soc.*, **70**, 644–647.
[1964i] Good ideals in fields of sets. *Ann. Math.*, **79**, 338–359.
[1964ii] Ultraproducts and saturated models. *Indag. Math.*, **26**, 178–186.
[1965] Limit ultraproducts. *J. Symb. Logic*, **30**, 212–234
[1965i] A survey of ultraproducts. *Logic, methodology and philosophy of science, proceedings of the 1964 international congress*, edited by Y. Bar-Hillel, North-Holland, Amsterdam.
[1967] Ultraproducts of finite sets. *J. Symb. Logic*, **32**, 47–57.
[1968] On the quantifier 'there exist uncountably many'. *Notices Am. Math. Soc.*, **15**, 654.
[1968i] Models with orderings. *Logic, methodology and philosphy of science* III, *proceedings of the third international congress, Amsterdam, 1967*, edited by B. van Rootselaar and J. F. Staal, North-Holland, Amsterdam, 35–62.
[1970] Logic with the quantifier 'there exist uncountably many', *Ann. Math. Logic*, **1** (1).
KELLEY, J. L.
[1955] *General topology*, Van Nostrand, Princeton.
KNEALE, W., and M. KNEALE
[1962] *The development of logic*, Clarendon Press, Oxford.
KOCHEN, S.
[1962] Ultraproducts in the theory of models. *Ann. Math.*, **74**, 221–261.
KREISEL, G.
[1956] Some uses of metamathematics. *Brit. J. Phil. Sci.*, **7**, 161–173.
[1967] Informal rigour and completeness proofs. *Problems in the philosophy of mathematics*, edited by I. Lakatos, North-Holland, Amsterdam, 138–186.
KREISEL, G., and J. L. KRIVINE
[1967] *Elements of mathematical logic: model theory*, North-Holland, Amsterdam.
KURATOWSKI, K. and A. MOSTOWSKI
[1967] *Set theory*, North-Holland, Amsterdam.
LEVY, A.
[1967] The sizes of the indescribable cardinals. *Proceedings of the 1967 Am. Math. Soc. summer institute on axiomatic set theory.*
ŁOŚ, J.
[1954] On the categoricity in power of elementary deductive systems and some related problems. *Colloq. Math.*, **3**, 58–62.

[1955] Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. *Mathematical interpretations of formal systems*, North-Holland, Amsterdam, 98–113.

LÖWENHEIM, L.

[1915] Über Möglichkeiten im Relativkalkül. *Math. Ann.*, **76**, 447–470. (English translation in VAN HEIJENOORT [1967], 228–251.)

MACDOWELL, R., and E. SPECKER

[1961] Modelle der Arithmetik. *Infinitistic methods, proceedings of the symposium on the foundations of mathematics, Warsaw, September 1959*, Pergamon Press, Oxford, 257–263.

MALCEV, A.

[1936] Untersuchungen aus dem Gebiete der mathematischen Logik. *Rec. Math. N.S.*, **1**, 323–336.

MATHIAS, A. R. D.

[1967] A survey of recent results in set theory, to appear in the *Proceedings of the 1967 Am. Math. Soc. summer institute on set theory.*

MENDELSON, E.

[1964] *Introduction to mathematical logic*, Van Nostrand, Princeton.

MOREL, A., D. SCOTT and A. TARSKI

[1958] Reduced products and the compactness theorem. *Notices Am. Math. Soc.*, **5**, 674–675.

MORLEY, M.

[1965] Categoricity in power. *Trans. Am. Math. Soc.*, **114**, 514–538.

[1968] Partitions and models. *Proceedings of the summer school in logic, Leeds, 1967*, edited by M. H. Löb, Springer, Berlin, 109–158.

MORLEY, M. and R. L. VAUGHT

[1962] Homogeneous universal models. *Math. Scand.*, **11**, 37–57.

MOSTOWSKI, A.

[1957] On a generalization of quantifiers. *Fund. Math.*, **44**, 12–36.

NEUMANN, B. H.

[1954] An essay on free products of groups with amalgamations. *Phil. Trans. Roy. London, Ser. A*, **246**, 503–554.

POST, E.

[1921] Introduction to a general theory of elementary propositions. *Am. J. Math.*, **43**, 163–185, (reprinted in VAN HEIJENOORT [1967], 264–284).

RABIN, M.

[1959] Arithmetical extensions with prescribed cardinality. *Indag. Math.*, **21**, 439–446.

RADO, R.

[1949] Axiomatic treatment of rank in infinite sets. *Can. J. Math.*, **1**, 337–343.

RASIOWA, H., and R. SIKORSKI

[1951] A proof of the completeness theorem of Gödel. *Fund. Math.*, **37**, 193–200.

[1963] *The mathematics of metamathematics*, Panstowe Wydawnictwo Naukowe, Warsaw.

ROBINSON, A.

[1956] *Complete theories*, North-Holland, Amsterdam.

[1963] *Introduction to model theory and to the metamathematics of algebra*, North-Holland, Amsterdam.

ROBINSON, A., and E. ZAKON

[1960] Elementary properties of ordered Abelian groups. *Trans. Am. Math. Soc.*, **96**, 222–236.

ROTMAN, B., and G. T. KNEEBONE

[1966] *The theory of sets and transfinite numbers*, Oldbourne, London.

RUBIN, H., and J. RUBIN

[1963] *Equivalents of the axiom of choice*, North-Holland, Amsterdam.

RUSSELL, B. A. W.

[1901] Recent works on the principles of mathematics. *The International Monthly*, **5**, 83–101 (reprinted as Mathematics and the metaphysicians in *Mysticism and logic*, (1918)).

SCOTT, D.

[1961] Measurable cardinals and constructible sets. *Bull. Acad. Polon. Sci., Sér. Sci. Math. Astr. Phys.*, **9**, 521–524.

[1961i] On constructing models for arithmetic. *Infinitistic methods, proceedings of the symposium on the foundations of mathematics*, Warsaw, September 1959, Pergamon Press, Oxford, 235–255.

SHOENFIELD, J.

[1961] The problem of predicativity. *Essays on the foundations of mathematics*, edited by Y. Bar-Hillel, E. I. J. Poznanski, M. O. Rabin and A. Robinson, North-Holland, Amsterdam, 132–139.

SIERPINSKI, W.

[1965] *Cardinal and ordinal numbers*, Panstowe Wydawnictwo Naukowe, Warsaw.

SIKORSKI, R.

[1964] *Boolean algebras*, 2nd ed., Springer, Berlin.

SILVER, J. H

[1966] *Some applications of model theory in set theory*. Doctoral dissertation, University of California at Berkeley.

SKOLEM, T.

[1920] *Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theorem über dichte Mengen*, Skrifter utgitt av videnskapsselskapit i Kristiania, I. Matematisk-naturvidenskabelig klasse, 4. (English translation in VAN HEIJENOORT [1967], 252–263.)

[1934] Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen. *Fund. Math.*, **23**, 150–161.

[1955] Peano's axioms and models of arithmetic. *Mathematical interpretations of formal systems*, North-Holland, Amsterdam, 1–14.

SLOMSON, A.

[1968] The monadic fragment of predicate calculus with the Chang quantifier and equality. *Proceedings of the summer school in logic, Leeds, 1967*, edited by M. H. Löb, Springer, Berlin, 279–302.

SOLOVAY, R. M.

[1967] A nonconstructible Δ^1_3-set of integers. *Trans. Am. Math. Soc.*, **127**, 50–75.

STEINITZ, E.

[1910] Algebraische Theorie der Körper. *J. Reine Angew. Math.*, **137**, 167–309.

STONE, M. H.

[1936] The representation theorem for Boolean Algebra. *Trans. Am. Math. Soc.*, **40**, 37–111.

[1937] Applications of the theory of Boolean rings to general topology. *Trans. Am. Math. Soc.*, **41**, 375–481.

TARSKI, A.

[1930] Une contribution à la théorie de la mesure. *Fund. Math.*, **15**, 42–50.

[1930i] Über einige fundamentale Begriffe der Metamathematik. *Compt. Rend. Séances Soc. Sci. Lett. Varsovie, cl. III*, **23**, 22–29. (English translation in TARSKI [1956], 30–37.)

[1935] Die Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica* (Warsaw), **1**, 261–405. (English translation in TARSKI [1956], 152–278.)

[1939] Ideale in vollständigen Mengenkörpern, *Fund. Math.*, **32**, 45–63.

[1949] Arithmetical classes and types of algebraically closed and real-closed fields. *Bull. Am. Math. Soc.*, **55**, 64.

[1952] Some notions and methods on the borderline of algebra and metamathematics. *Proceedings of the international congress of mathematicians, 1950*, vol. I, 705–720.

[1954–5] Contributions to the theory of models, I, II, III. *Indag. Math.*, **16**, 572–588, **17**, 56–64.

[1956] *Logic, semantics, metamathematics, papers from 1923 to 1938*, translated by J. H. Woodger, Clarendon Press, Oxford.

[1958] Remarks on predicate languages with infinitely long expressions. *Colloq. Math.*, **6**, 171–176.

[1962] Some problems and results relevant to the foundations of set theory. *Logic, methodology and philosophy of science, proceedings of the 1960 international congress*, edited by E. Nagel, P. Suppes and A. Tarski, Stanford U.P., 125–135.

TARSKI, A., and R. L. VAUGHT

[1957] Arithmetical extensions of relational systems. *Compositio Math.*, **13**, 81–102.

ULAM, S.

[1930] Zur Masstheorie in der allgemeinen Mengenlehre. *Fund. Math.*, **16**, 140–150.

VAUGHT, R. L.

[1954] Applications of the Löwenheim–Skolem–Tarski theorem to problems of completeness and decidability. *Indag. Math.*, **16**, 467–472.

[1956] On the axiom of choice and some metamathematical theorems. *Bull. Am. Math. Soc.*, **62**, 262–263.

[1963] Models of complete theories. *Ibid.*, **69**, 299–313.

[1964] The completeness of logic with the added quantifier 'there are uncountably many'. *Fund. Math.*, **54**, 303–304.

VAN DER WAERDEN, B. L.

[1953] *Modern algebra, vol. I*, Ungar, New York.

WILMERS, G.

[1967] *Saturated systems*. Diploma dissertation, Oxford.

YASUHARA, M.

[1966] An axiomatic system for the first order language with an equi-cardinality quantifier. *J. Symb. Logic*, **31**, 635–640.

INDEX OF NAMED THEOREMS

Beth's theorem	173.
Chang's two cardinal theorem	259
Compactness theorem for L_Q, The	279
Compactness theorem for predicate calculus, The	102
Compactness theorem for propositional calculus, The	45
Completeness theorem for L_Q, The	281
Completeness theorem for predicate calculus, The	62
Completeness theorem for propositional calculus, The	43
Consistency of predicate calculus, The	60
Craig's interpolation lemma	156, 171
Deduction theorem for predicate calculus, The	66
Deduction theorem for propositional calculus, The	43
Downward Löwenheim–Skolem theorem, The	81
Downward Löwenheim–Skolem theorem for L_{Q_α}, The	266
Finiteness theorem for predicate calculus, The	60
Finiteness theorem for propositional calculus, The	39
Frayne's lemma	161
Fuhrken's compactness theorem	277
Fuhrken's first normal form theorem	271
Fuhrken's second normal form theorem	275
Generalized completeness theorem for propositional calculus, The	47
Gödel's completeness theorem	234
Gödel–Henkin completeness theorem, The	102
Hanf's incompactness theorem	293
Hanf's theorem	85
Herbrand's theorem	236
Homogenization theorem, The	211
Isomorphism theorem, The	207
Jónsson's theorem	213
Keisler's ultrapower theorem	150
Kochen's ultralimit theorem	168
Lindenbaum's theorem	182
Łoš's theorem	90
Łoš's theorem for $L_{\kappa\kappa}$	296
Łoš's theorem for L_Q	268
Łoš's theorem for L_{Q_α}	263

Löwenheim–Skolem theorem, The 82
MacDowell and Specker's theorem 244
Prime model test, The 184
Rabin–Keisler theorem, The 139
Regular compactness theorem for L_Q, The 276
Robinson's model completeness test 191
Scott's lemma 163
Scott's theorem 305
Singular compactness theorem for L_Q, The 278
Skolem's normal form theorem 233
Stone representation theorem, The 24
Tarski's lemma 21
Ultrafilter theorem, The 15
Union of chains 79
Upward Löwenheim–Skolem theorem, The 82
Vaught's completeness theorem for L_{Q_1} 280
Vaught's test 179
Vaught's test for L_{Q_α} 266
Vaught's two cardinal theorem 251

GENERAL INDEX

adherent 22
adherent point 22
admits the pair of cardinals 246
algebraically closed fields 99, 197, 200
— of specified characteristic 177, 179
α-saturated 218
amalgamation property 203
antisymmetry 7
appropriate language 55
atom 27
atomic Boolean algebra 27
atomic formulas 51, 68, 70
— of $L_{\alpha\beta}$ 287
atomless Boolean algebra 28, 180
axiom of choice 4
axiom of constructibility 301
axioms 36

base 14
basic symbols
— of SC 32
— of PC 50
β-incompact 289
beth number 6, 205
beths 6, 205
Boolean algebra 11
— atomic 27
— atomless 28, 180
— dense 30
Boolean space 24
bound variable 53

canonical embedding 92
canonical homomorphism 19
Cantor discontinuum 30
Cantor space 28
Cantor's ternary set 30
cardinal 5
— compact 290
— inaccessible 5
— incompact 290
— limit 5
— measurable 112, 295
— of a language 80
— ω-measurable 112
— regular 5
— singular 5
— strong limit 5
— strongly compact 290
— strongly inaccessible 5
— successor 5
— weakly compact 290
cardinal like
— model 244
— ordering 244
categorical 175
chain 8
— elementary 79
— of relational structures 79
Chang quantifier 268
characteristic algebra 12, 25
characteristic p 96
characteristic zero 96
classes 300
clopen 12, 24
closure 59
cofinal 5
cofinality 5
cofinite 11
commutative fields 96
commutative groups 95
commutative rings 95

compact cardinal 290
compactness property 289
complement 10
complemented lattice 10
complete axiomatization 42
complete lattice 9
complete set of connectives 36
complete theory 178
component 24
connected 24, 49
consistent 44, 45, 60
constants 68
constant symbols 68
continuous chain 205
converge 22
cover existentially 186
cover universally 186

deducible 60
definable function 238
degree 50
Δ-elementary class 141
$\Delta\Sigma$-elementary class 141
dense Boolean algebra 30
densely ordered sets 94, 176, 179, 191–194
describability 305
diagram 183
dichotomy 7
discretely ordered sets 194–197
disjointed function 117
distributive lattice 10
divisible group 180
domain 55
duality 13

elementary chain 79
elementary class 141
elementary embedding 75
elementary extension 75
elementary substructure 75
elementarily embeddable 75
elementarily equivalent 74
embedding 73
end extension 240
enumeration 78
— without repetitions 78
equality symbol 51
equi-cardinal quantifier 268
existential formula 185
existential substructure 186
expansion 153, 154
explicitly definable 172
extension 73
— elementary 75
— end 240
— proper 240, 297

false 34
fields 96
— algebraically closed 99, 197–200
filter 12
— neighbourhood 22
— prime 20
— principal 12
— proper 13
finite dimensional vector space 98
finite intersection property 13
finitely axiomatizable 92
finitely satisfiable 218
fip 13
first order property 92
formal arithmetic 237
formula 33, 51, 68, 70
— atomic 51, 68, 70
— existential 185
— open 59, 185
— primitive 190
— provable 37, 58
— quantifier free 185
— universal 185
— valid 35
free variable 53
free for 54, 70
full expansion 214
function letters 70

gags 3, 6, 48, 158
G.C.H. 6
general first order property 93
Generalization 58
generalized continuum hypothesis 6
generalized quantifier 261
generate 14
good ultrafilter 117

graph 49
greatest lower bound 8
groups 95
— commutative 95
— divisible 180
— torsion free 180

Hanf number 85
homogeneous 226
homomorphism 17, 73
— canonical 19

ideal 12
— principal 12
identity symbol 51
iff 6
immediate consequence 37
implicitly definable 172, 173
inaccessible cardinal 5
incompact cardinal 290
individual variables 50
induced realization 43
infimum 8
initial segment 244
injection map 73
irrefutable sentence 64
isomorphic 18
isomorphic ultrafilter pairs 116
isomorphism 17, 73

join 8

κ-class 203
k-chromatic graph 49
kernel 20
Kolmogoroff space 158

$L_{\alpha\beta}$-structure 288
lattice 8
— complemented 10
— complete 9
— distributive 10
— $\mathfrak{m}$-complete 9
L.C.H. 5, 278
least upper bound 8
less than κ-complete 295
$L_{\kappa\kappa}$-equivalent 296
$L_{\kappa\kappa}$-substructure 296
limit beth number 205
limit cardinal 5
limit cardinal hypothesis 5, 278
limit ordinal 5
Lindenbaum algebra 42, 61
locally consistent 289
logical connectives 32, 51
logical symbol 51
Löwenheim number 85
lower bound 8

maximal 9
maximum 9
$\mathfrak{m}$-categorical 176
$\mathfrak{m}$-compact 263
$\mathfrak{m}$-complete lattice 9
measurable cardinal 112, 295
meet 8
$\mathcal{M}$-homogeneous 202
$\mathcal{M}$-homoiogeneous 207
minimum 9
$(\mathcal{M}, \kappa)$-homogeneous 202
$(\mathcal{M}, \kappa)$-universal 202
$\mathfrak{m}$-like model 244
model 56
model complete 183
monotonic function 117
Modus Ponens 37, 58
$\mathcal{M}$-special 213
multiplicative closure 117
multiplicative function 117
multiplicative set 117
$\mathcal{M}$-universal 202
μ-structure 288

neighbourhood filter 22
$\mathfrak{n}$-normal 268
non-principal ultrafilter 15, 108
non-standard model of arithmetic 237
normal vector space 248

ω-complete ultrafilter 111
ω-measurable cardinal 112
open formula 59, 185
order 70

ordinal 4
— limit 5
— successor 5

parametrically definable function 242
partially ordered set 7
partition 113
perfect 28
Pow 151
power set 6, 7
power set Boolean algebra 11
predicate letter 50
prenex normal form 185
preserve meets 21
prime filter 20
prime model 184
primitive formula 190
principal filter 12
principal ideal 12
principal ultrafilter 15, 107
Prod 151
projection 167
proof 37, 58
proof from assumptions 39
proper extension 240, 297
proper filter 13
propositional variable 32
provable formula 37, 58
provable from assumptions 59
pseudo-elementary class 154
punctuation symbols 33, 51

Q_α 262
quantifier free formula 185
quantifier symbol 51
quotient algebra 19

realization 34, 55, 68, 70
— induced 43
— of $L_{\alpha\beta}$ 288
reduced power 89
reduced product 89
reduct 153
reflexivity 7
refutable sentence 64
regular cardinal 5
regular ultrafilter 114
relational structure 55
— countable 64
— of cardinal α 64
relativization 249
relativizing quantifiers 181, 249
restriction 73
rings 75
rules of inference 36

satisfaction 68, 70
satisfiable 35, 218
satisfies 35, 56, 289
SC 32
scope of a quantifier 52
sentence 53
— irrefutable 57
— refutable 57
— true 56
— universally valid 56
— valid 56
sets 300
shell 20
Σ-elementary class 141
$\Sigma\Delta$-elementary class 141
singular cardinal 5
Skolem normal form 233
small for 263
special 230
specializing chain 230
standard model of arithmetic 237
Stone space 26
strict ordering 8
string 33, 51
strong limit cardinal 5
strongly incompact 290
strongly inaccessible cardinal 5
subalgebra 17
subbase 14
subrealization 73
substructure 73
successor cardinal 5
successor ordinal 5
supremum 8

τ-acceptable 288
tautology 35
term 70

theory 175
— categorical 175
— complete 178
— model complete 183
theory of $\mathfrak{A}$ 140
theory of complete arithmetic 237
theory of formal arithmetic 237
torsion free group 180
totally disconnected 24
totally ordered set 7
transitive class 301
transitivity 7
true 34, 56
type 73

ultrafilter 14
— α-complete 111
— good 117
— non-principal 15, 108
— ω-complete 111
— principal 15, 107
— regular 114
— uniform 111
ultrafilter on 22, 87, 107
ultrafilter pair 107
— isomorphic 116
ultrafilter sequence 165
ultralimit 165, 167
ultrapower 89
ultraproduct 89
uniform ultrafilter 111
universal 226
universal formula 185
universal substructure 186
universally valid 56
upper bound 8

vacuous 53
valid 56
valid formula 35
valuation 55
variables 50
— bound 53
— free 53
vector spaces 97
— finite dimensional 98, 181

weakly compact 290
well founded 135
well ordering 133
well ordering principle 4

x-quantifier 52

Zorn's lemma 4, 9

STUDIES IN LOGIC

AND THE FOUNDATIONS OF MATHEMATICS

W. ACKERMANN †	*Solvable Cases of the Decision Problems • (1968)*
J. W. ADDISON, L. HENKIN, A. TARSKI (Editors)	*The Theory of Models • (1972)*
Y. BAR-HILLEL (Editor)	*Logic, Methodology and Philosophy of Science • (1972)*
Y. BAR-HILLEL (Editor)	*Mathematical Logic and the Foundations of Set Theory • (1970)*
P. BERNAYS, A. A. FRAENKEL †	*Axiomatic Set Theory • (1968)*
E. W. BETH †	*The Foundations of Mathematics • (1965)*
W. W. BOONE, F. B. CANNONITO, R. C. LYNDON (Editors)	*Word Problems • (1973)*
L. BORKOWSKI (Editor)	*Jan Łukasiewicz, Selected Works • (1970)*
C. C. CHANG, H. J. KEISLER	*Model Theory • (1973)*
W. CRAIG	*Logic in Algebraic Form • (1974)*
H. B. CURRY	*Outlines of a Formalist Philosophy of Mathematics • (1970)*
H. B. CURRY, R. FEYS †	*Combinatory Logic, Volume I • (1974)*
H. B. CURRY, J. R. HINDLEY, J. P. SELDIN	*Combinatory Logic, Volume II • (1972)*
F. R. DRAKE	*Set Theory: An Introduction to Large Cardinals • (1974)*
J. E. FENSTAD (Editor)	*Proceedings of the Second Scandinavian Logic Symposium • (1971)*
J. E. FENSTAD, P. G. HINMAN (Editors)	*Generalized Recursion Theory • (1974)*
M. C. FITTING	*Intuitionistic Logic, Model Theory and Forcing • (1969)*
A. A. FRAENKEL †	*Abstract Set Theory • (1968)*
A. A. FRAENKEL †, Y. BAR-HILLEL, A. LEVY	*Foundations of Set Theory • (1973)*
R. O. GANDY, C. E. M. YATES (Editors)	*Logic Colloquium '69 • (1971)*
L. HENKIN, J. D. MONK, A. TARSKI	*Cylindric Algebras, Part I • (1971)*
A. HEYTING	*Intuitionism • (1971)*
T. J. JECH	*The Axiom of Choice • (1973)*
H. J. KEISLER	*Model Theory for Infinitary Logic • (1971)*
A. KINO, J. MYHILL, R. E. VESLEY (Editors)	*Intuitionism and Proof Theory • (1970)*
G. KREISEL, J. L. KRIVINE	*Elements of Mathematical Logic • (1971)*
K. KURATOWSKI, A. MOSTOWSKI	*Set Theory • (1968)*
I. LAKATOS (Editor)	*Problems in the Philosophy of Mathematics • (1972)*
I. LAKATOS (Editor)	*The Problem of Inductive Logic • (1968)*
I. LAKATOS, A. MUSGRAVE (Editors)	*Problems in the Philosophy of Science • (1968)*
H. LEBLANC (Editor)	*Truth, Syntax and Modality • (1973)*
W. A. J. LUXEMBURG, A. ROBINSON (Editors)	*Contributions to Non-Standard Analysis • (1972)*
A. I. MAL'CEV †	*The Metamathematics of Algebraic Systems • (1971)*
Y. N. MOSCHOVAKIS	*Elementary Induction on Abstract Structures • (1974)*
A. MOSTOWSKI	*Sentences Undecidable in Formalized Arithmetic • (1964)*
A. MOSTOWSKI	*Constructible Sets with Applications • (1969)*
H. RASIOWA	*An Algebraic Approach to Non-Classical Logics • (1974)*
A. ROBINSON	*Non-Standard Analysis • (1974)*
H. RUBIN, J. E. RUBIN	*Equivalents of the Axiom of Choice • (1970)*
Th. SKOLEM † et al.	*Mathematical Interpretation of Formal Systems • (1971)*
M. E. SZABO (Editor)	*The Collected Papers of Gerhard Gentzen • (1969)*
A. TARSKI et al.	*Ordinal Algebras • (1970)*
A. TARSKI, A. MOSTOWSKI, R. M. ROBINSON	*Undecidable Theories • (1971)*
P. VOPĚNKA, P. HÁJEK	*The Theory of Semisets • (1972)*

AMSTERDAM P.O. BOX 211 **NORTH HOLLAND** PUBLISHING CO. NETHERLANDS

BELL/SLOMSON MODELS AND ULTRAPRODUCTS

511.
3
BEL